Structural Modelling of Small Molecules by NMR: Solution-State Structure of 12,13-Diepimeric Coenzyme F430 and Comparison with the X-ray Structure of the Pentamethyl Ester Derivative

Hoshik Won,[†] Karl D. Olson,[‡] Dennis R. Hare,[#] Ralph S. Wolfe,*[,‡] Christoph Kratky,*[,§] and Michael F. Summers*[,†]

Contribution from the Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, Maryland 21228, Department of Microbiology, University of Illinois—Urbana Champaign, Urbana, Illinois 61801, Hare Research Inc., 14810 216th Avenue N.E., Woodinville, Washington 98072, and Institut für physikalische Chemie der Universität Graz, Graz, Austria. *Received December 23, 1991*

Abstract: Two-dimensional nuclear magnetic resonance (2D NMR) spectroscopic methods have been employed in combination with distance geometry and Overhauser effect back-calculation computational methods to determine the solution-state structure of the 12,13-diepimeric form of coenzyme F430 (12,13-diepimeric F430). NMR signal assignments were made for all non-exchangeable protons using double quantum filtered correlated (DQF-COSY), homonuclear Hartmann-Hahn (HOHAHA), nuclear Overhauser effect (NOESY), and rotating frame Overhauser effect (ROESY) ^1H-^1H correlated spectroscopies. Assignments were made for most carbon atoms via heteronuclear multiple quantum coherence (HMQC) and heteronuclear multiple bond (HMBC) ^1H-^{13}C correlated spectroscopies. The three-dimensional structure of 12,13-diepimeric F430 was determined with an approach developed originally for studies of biopolymers that includes back-calculation of 2D nuclear Overhauer effect spectra for model structures generated by DG/SA computations. As a first step, primary restraints were derived that allowed for maximum conformational flexibility in the absence of NOE-derived distance information. Structures generated subsequently with "loose" NOE-derived distance restraints (2.0–2.5, 2.0–3.5, and 2.0–4.5 Å for strong, medium, and weak NOE cross peak intensities, respectively) converged to a unique conformation, with pairwise root-mean-square deviation (RMSD) values for the macrocycle atoms in the range 0.019 to 0.153 Å. Except for some side-chain atoms, the back-calculated NOESY spectra of the DG/SA structures were consistent with the experimental NOESY spectrum. Most of the side-chain groups did not converge to unique conformations, and pairwise RMSD values for all atoms were in the range 0.67 to 1.25 Å. The NMR structure is compared with the recently determined X-ray structure of the pentamethylated ester derivative, 12,13-diepimeric F430M. Both structures exhibit a very similar corphin ring conformation that is characterized by an extreme saddle-shaped deformation. The average RMS difference between the 24 corphin ring atoms of the X-ray structure and the NMR structures is 0.28 Å. Larger differences are observed only for atoms of the flexible side chains. The similarity of the NMR and X-ray structures provides grounds for optimism that the NMR approach may be extendable in a general way to studies of smaller molecules that have proven difficult to crystallize, including the *native* form of coenzme F430.

Introduction

Coenzyme F430 is a nickel(II)-containing cofactor for methyl coenzyme M reductase that is utilized by methanogenic bacteria.[1-4] By an unknown mechanism, F430 mediates the reductive demethylation of methyl coenzyme M (2-(methylthio)ethanesulfonic acid, CH_3-S-CoM) using reducing equivalents from (7-mercaptoheptanoyl)threonine phosphate (HS-HTP). This is the last step in the conversion of CO_2 into methane in what is believed to be an energy-yielding process in methanogens.[5,6] The corphin ring of F430, a term coined by Eschenmoser and co-workers to indicate that is a structural hybrid of porphyrins and corrins,[7] is the most highly reduced tetrapyrrole known to exist biologically and is one of only two nickel-containing macrocycles known to occur in the biological world.

To date, the atomic-level structural characterization of native coenzyme F430 has not been achieved due, in part, to difficulties in obtaining crystals suitable for X-ray diffraction studies. The molecular constitution and most of the configurational structure have been established by a series of chemical and NMR spectroscopic studies of the pentamethyl ester derivative, F430M.[7-10] In addition, the pentamethyl ester derivative of the 12,13-diepimeric form of F430 (12,13-diepi-F430M; the 12,13-diepimer is thermodynamically more stable than the native coenzyme) has been crystallized, and the atomic-level structure of this thermodynamically more stable form of the coenzyme has recently been determined from an X-ray crystal structure analysis.[11] In combination with chemical and NMR evidence, this structure analysis allowed assignment of the configuration of native F430, reversing the original stereochemical assignments for the C17, C18, and C19 heterocyclic atoms.[8]

Three-dimensional structural analysis of the *native* form of F430 is essential for rationalizing spectroscopic data obtained for the holoenzyme and for evaluating potential mechanistic pathways.[3,12-14] Structural predictions have thus far been based on

[†] University of Maryland.
[‡] University of Illinois.
[#] Hare Research, Inc.
[§] Universität Graz.

(1) Gunsalus, R. P.; Wolfe, R. S. *FEMS Microbiol. Lett.* **1978**, *3*, 191.
(2) Diekert, G.; Konheiser, U.; Thauer, R. K. *J. Bacteriol.* **1981**, *148*, 459.
(3) Walsh, C. T.; Orme-Johnson, W. H. *Biochemistry* **1987**, *26*, 4901.
(4) Rouviere, P. E.; Wolfe, R. S. *J. Biol. Chem.* **1988**, *263*, 7913.
(5) Bobik, T. A.; Olson, K. D.; Noll, K. M.; Wolfe, R. S. *Biochem. Biophys. Res. Commun.* **1987**, *149*, 455.
(6) Ellermann, J.; Hedderich, R.; Böcher, R.; Thauer, R. K. *Eur. J. Biochem*, **1988**, *172*, 669.
(7) Fässler, A.; Pfaltz, A.; Mueller, P. M.; Farooq, S.; Kratky, C.; Kraeutler, B.; Eschenmoser, A. *Helv. Chim. Acta* **1982**, *65*, 812.
(8) Pfaltz, A.; Jaun, B.; Faessler, A.; Eschenmoser, A.; Jaenchen, R.; Gilles, H. H.; Diekert, G.; Thauer, R. K. *Helv. Chim. Acta* **1982**, *65*, 828.
(9) Livingston, D. A.; Pfaltz, A.; Schreiber, J.; Eschenmoser, A.; Ankel-Fuchs, D.; Moll, J.; Jaenchen, R.; Thauer, R. K. *Helv. Chim. Acta* **1984**, *67*, 334.
(10) Pfaltz, A.; Livingston, D. A.; Jaun, B.; Diekert, G.; Thauer, R. K.; Eschenmoser, A. *Helv. Chim. Acta* **1985**, *68*, 1338.
(11) Färber, G.; Keller, E.; Kratky, C.; Jaun, B.; Pfaltz, A.; Spinner, C.; Kobelt, A.; Eschenmoser, A. *Helv. Chim. Acta* **1991**, *74*, 697.
(12) Hamilton, C. L.; Scott, R. A.; Johnson, M. K. *J. Biol. Chem.* **1989**, *264*, 11605.
(13) Pfaltz, A. In *The Biochemistry of Nickel*; Lancaster, J. R., Jr., Ed.; VCH Publishers: New York, 1988; pp 275–298.
(14) Eschenmoser, A. *Ann. N. Y. Acad. Sci.* **1986**, *471*, 108.

0002-7863/92/1514-6880$03.00/0 © 1992 American Chemical Society

molecular mechanics (MM2) calculations, where a trigonal bipyramidal Ni(II) coordination geometry has been proposed for the holoenzyme,[15] and on analyses of crystal structure data obtained for a series of high-spin and low-spin hydrocorphinoid Ni(II) model complexes.[16-19] According to the latter analysis,[14,18] the dominant response to the contraction of the central N_4-coordination hole to accommodate a "small" metal center (such as diamagnetic Ni(II)) is a saddle-shaped overall conformation of the corphin ring. The extent of saddle deformation depends on the molecular constitution of the porphinoid ligand (highly saturated porphyrin rings exhibit larger saddle deformations) and on the metal atom's ability to induce a contraction of the hydrocorphinoid ring as measured, for example, by the transannular N–N distances. The quantity d_m, defined as the mean absolute deviation of the four meso-C atoms from a least-squares plane through the four pyrrolic N atoms ($d_m = (|dC5| + |dC10| + |dC15| + |dC20|)/4$), was introduced as a quantitative measure for the extent of saddle deformation, and was shown to correlate inversely with the metal–N and the transannular N–N distances.[11,18] The relationship between the saddle-shaped deformation and the conformations of the hydrogenated pyrrolic rings has been discussed in terms of *conformation coupling*. This phenomenon yields, among other results, an explanation for the higher thermodynamic stability of 12,13-diepi-F430 compared to native F430, as well as its lower tendency toward conversion in the high-spin form, and it suggests that 12,13-diepi-F430 is less strained than native F430.[11,18]

Given the absence of suitable crystals for X-ray diffraction studies, NMR spectroscopy provides an alternative approach for establishing the structural features of native F430. However, methodologies developed for NMR-based structure determination of biopolymers[20,21] have not been extended in a general manner to smaller molecules. A potentially serious problem results from the fact that small molecules typically do not contain the "long-range" dipolar interactions observed for polymers (i.e., interactions involving protons that are close in space but are separated by one or more subunits of the polymer). In proteins, for example, proton dipolar interactions are generally observed among nonnearest neighbor residues, such as those on opposite strands of a β-sheet, allowing backbone foldings to be established even when only loose distance restraints are employed during structure refinement.[20,21]

The availability of high-resolution X-ray structural data for 12,13-diepi-F430M[11] has provided us with the opportunity to carry out NMR-based 3D structural modelling with the objective of comparing the results of the two approaches. If the NMR-derived structures compared favorably with the X-ray structures, the approach might be extendable to the structure determination of the thus-far elusive native F430, and perhaps also to other small molecules that have proven difficult to crystallize. We have utilized a relatively new NMR-based structure determination approach that involves the application of loose distance restraints to represent the NOE data in hybrid distance geometry/simulated annealing (DG/SA) modelling (as is commonly used for protein structure determination), and includes 2D NOE back-calculations for establishing the consistence between the NMR-derived models and the experimental data.[22-24] The results of the NMR-based structural modelling the comparisons with the X-ray structure are presented, and the potential and limitations of the NMR approach are discussed.

Experimental Section

Materials. Phenyl Sepharose CL-4B, QAE A-25, and DEAE-Sephadex A-25 were purchased from Pharmacia LKB Biotechnology Inc. C18 PR-HPLC (reverse-phase high pressure liquid chromatography) columns were obtained from Waters. Deuterated trifluoroethanol was from Cambridge Isotopes. Bacterial cells were lab stock cultures.

Bacterial Cells. *Methanobacterium thermoautrophicum* strain ΔH (DMS 1053) as grown in a 250-L fermenter (B. Braun) at 60 °C, pH = 7.3, with H_2 and CO_2 as energy and carbon sources, respectively. The medium was reduced with H_2S (to ca. −440 mV versus NHE) before inoculation. During the fermenter run, H_2S, H_2, and CO_2 flow rates were adjusted manually and via computer to maintain a constant pH (±0.15). As the cells reached the end of exponential growth (but before stationary phase), they were harvested anaerobically (Sharples centrifuge). The cell paste was then immediately transferred into Wheaton bottles and made anaerobic by several nitrogen gas flushing cycles in an air lock of an anaerobic hood (Coy). The cells were stored under N_2 at −20 °C as either a cell paste or as a cell suspension of whole cells in 50 mM potassium phosphate (pH 7.0) buffer (1:1).

Isolation and Purification of F430. Stereoisomers of F430 (native, 12,13-diepi-F430, and 13-monoepi-F430 as well as the oxidized analogue, 12,13-didehydro-F430 (F560)) were purified using hydrophobic interaction chromatography (Phenyl Sepharose) and anion exchange chromatography (QAE A-52 and DEAE A-25) as described previously.[25,26] All steps, except RP-HPLC (reverse phase), were performed at 5 °C.

12,13-Diepi-F430 was further purified with an additional RP-HPLC system. A Waters Model 501 solvent programmer was used to run a nonlinear gradient solvent system on Waters Model 501 pumps. The pressure was held below 3000 psi. The detector was set to read absorbances at 431 nm or 214 nm (Waters variable UV detector). The progress of the gradient was monitored by a Hewlett-Packard chart recorder. The nonlinear gradient was run under the following conditions: Waters C_{18} RP-HPLC column (3.9 mm × 30 cm), 25-min nonlinear gradient, 50 mM NH_4CO_2H, pH 7.0 (10% MeOH) to 50 mM NH_4CO_2H, pH 7.0 (50% MeOH), 0.5 mL/min, detector (431 nm and 214 nm). The diepimer F430 eluted off the column at t_R = 23.5 min (N = 2800). The effluents were pooled and lyophilized three times to remove the ammonium formate. The final 12,13-diepi-F430 has a characteristic UV/vis spectrum with A430/A275 (A = absorbance) = 1.17 (in contrast of 1.1 for the native coenzyme). The 12,13-diepi-F430 also exhibited shoulders at 300 nm and 340 nm and the width at one-half height of the 430-nm peak was 42 nm (versus 62 nm for the native coenzyme) (see supplementary Figure S1).

The FAB (fast atom bombardment) cation mass spectrum of 12,13-diepi-F430 gave an expected peak at m/z = 905 (matrix = nitrobenzyl alcohol). The elemental composition was determined to be $C_{42}H_{51}O_{13}N_6Ni$ by high resolution mass spectrometry. The 70-SE-4F (70-4F) is a high resolution, four-sector 8-kV mass spectrometer.

NMR Data. NMR spectral data (500.14 MHz, ^1H) were obtained using a GE GN-500 spectrometer. Sample conditions were as follows: 1.13 mM 12,13-diepi-F430 in F_3CCD_2OD (TFE-d_3) solvent; T = 25 °C. Raw NMR data were transferred via ethernet to Silicon Graphics Personal Iris computers, converted to "readable" files using an in-house program (GENET) and processed with FTNMR (Hare Research, Inc.). ^1H and ^{13}C NMR chemical shifts were referenced to internal TFE-d_3 (3.88 ppm, ^1H; 61.5 ppm, ^{13}C). Homonuclear correlated spectra were zero-filled to final matrix sizes of 1024 × 1024 real points. Data acquisition and processing parameters for individual 2D experiments (defined below) were as follows. Double quantum-filtered correlated spectroscopy,[27] **2QF-COSY**: 96 scans per t_1 increment; 2.8-s repetition delay; 2 × 256 × 1024 raw data matrix size, processed with 4-Hz Gaussian filtering in the t_2 dimension and 8-Hz Gaussian plus trapezoidal filtering in the t_1 dimension. Homonuclear Hartmann–Hahn spectroscopy,[28-31] **HOHAHA**: 2 × 256 × 1024 raw data matrix size; 32 scans per t_1

(15) Zimmer, M.; Crabtree, R. H. *J. Am. Chem. Soc.* **1990**, *112*, 1062.
(16) Kratky, C.; Angst, C.; Johnasen, J. E. *Angew. Chem., Int. Ed. Engl.* **1981**, *20*, 211.
(17) Kratky, C.; Fässler, A.; Pfaltz, A.; Kräutler, B.; Juan, B.; Eschenmoser, A. *J. Chem. Soc., Chem. Commun.* **1984**, 1368.
(18) Kratky, C.; Waditschatka, R.; Angst, C.; Jonasen, J. E.; Plaquwvent, J. C.; Schreiber, J.; Eschenmoser, A. *Helv. Chim. Acta* **1985**, *68*, 1312.
(19) Waditschatka, R.; Kratky, C.; Jaun, B.; Heinzer, J.; Eschenmoser, A. *J. Chem. Soc., Chem. Commun.* **1985**, 1604.
(20) Wurtrich, K. *Science* **1989**, *243*, 45.
(21) Clore, G. M.; Gronenborn, A. M. *Science* **1991**, *252*, 1390.
(22) Summers, M. F.; South, T. L.; Kim, B.; Hare, D. R. *Biochemistry* **1990**, *29*, 329.
(23) Blake, P.; Hare, D. R.; Summers, M. F. *Techniques in Protein Chemistry*; Villefranca, J. J., Ed.; Academic Press: New York, 1991; p 357.
(24) South, T.; Blake, P.; Hare, D.; Summers, M. F. *Biochemistry* **1991**, *30*, 6342.
(25) Won, H.; Olson, K. D.; Wolfe, R. S.; Summers, M. F. *J. Am. Chem. Soc.* **1990**, *112*, 2178.
(26) Shiemke, A. K.; Shelnutt, J. A.; Scott, R. A. *J. Biol. Chem.* **1989**, *64*, 11236.
(27) Rance, M.; Sorensen, O. W.; Bodenhausen, B.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *17*, 479.
(28) Braunschweiler, L.; Ernst, R. R., *J. Magn. Reson.* **1983**, *53*, 521.
(29) Davis, D. G.; Bax, A. *J. Am. Chem. Soc.* **1985**, *107*, 2820.
(30) Davis, D. G.; Bax, A. *J. Am. Chem. Soc.* **1985**, *107*, 7197.
(31) Bax, A.; Davis, D. G. In *Advanced Magnetic Resonance Techniques in Systems of High Molecular Complexity*; Nicolai, N., Valensin, G., Eds.; Birkhauser: Boston, **1986**; p 21.

Table I. Scalar and Dipolar NMR Connectivities Used to Assign ^1H and ^{13}C Signal of 12,13-Diepi-F430[a]

^1H signal	HOHAHA	COSY	NOESY	HMBC
Me2	none	none	H2a (s), H2a′ (s), H3 (w), H3a (s), H3a′ (s) H3b,b′ (w), H5 (s), H5′ (m), H20,20′ (s)	C1, C2, C2a, C3
Me7	none	none	H5 (s), H8 (w), H8a′ (m), H8b′ (s), H8b (s) H8a (m, overlaps with H5′)	C6, C7, C7a, C8
H2a	H2a′	H2a′	Me2 (s), H2a′ (s)	C1, C2, C2b, C3
H2a′	H2a	H2a	H2a (s), Me2 (s), H20,20′ (s)	C2, C2b, C3
H3	H3b,b′, H5 H5′, H3a, H3a′	H3a, H3a′	Me2 (s), H3a (s), H3a′ (s) H3b,b′ (s, overlaps with H12a,a′)	C3a
H3a	H3a′, H5	H3a′	H3a′ (s), H5 (s), Me2 (s)	C3c
H3a′	H3a, H5	H3a	H3a (s), H5 (w), Me2 (s)	C4
H3b,b′	H3a, H3a′ H5, H5′	H3a, H3a′	Me2 (w), H3a (s), H5 (w) H3a′ (s, overlaps with H17a)	C3, C3a, C3c
H4	H3, H3b,b′, H5′ H3a, H3a′, H5	H3, H5	H3 (s), H3b,b′ (m) H5 (m), H5′ (s)	C1, C2
H5	H5′	H5′	Me2 (s), H5′ (s), Me7 (s)	C4, C6
H5′	H3a, H3a′, H5	H5	Me2 (m), H5 (s), H3a (s), H3a′ (s)	C6, C4
H7a,a′	none	none	Me7 (s) (H8b)	C6, C7, C7b
H8	H8a, H8a′ H8b, H8b′	H8a	Me7 (w), H8a (s), H8b (s), H8b′ (s) H8a′ (s, overlaps with H12a,a′)	C7a, C8a, C9
H8a	H8a′	H8a′	H8a′ (s), Me7 (m)	C8c
H8a′	H8a	H8a, H8b	H8a (s), Me7 (m)	C8c
H8b	H8a, H8a′, H8b′	H8a, H8b′	Me7 (s), H8b′ (s) H8a (s, overlaps with H13a) H8a′ (s, overlaps with H12a,a′)	C8, C8c
H8b′	H8a′, H8a, H8b	H8b	H8a (s, overlaps with H13a), H8b (s), H8b′ (s)	C8, C8c
H10	none	none	H12 (s), H8 (s), H8b′ (w), H8b (w), H8a (s) H8a′ (s, overlaps with H12a,a′)	C8, C9, C12, C11
H12	H12a,a′	H12a,a′	H13b′ (w), H13b (m), H13a′ (m) H12a,a′ (s, overlaps with H13a′), H13a (m)	C11, C12a, C12b, 13a
H12a,a′	H13a, H13a′	H12	H13a (s, overlaps with H8a)	C11, C12b, C13
H13	H13a, H13a′ H13b, H13b′	H13a	H12 (s), H13b′ (m), H13b (m) H12a,a′ (s, overlaps with H13a′), H13a (s)	C11, C12a, C13a
H13a	H13a′	H13a′	H13a′ (s)	C13, C13c
H13a′	H13a	H13a	H13a (s, overlaps with H8a)	C13
H13b	H13a, H13a′, H13b′		H13a′, H13b′, H13a (s, overlaps with H8a) H13a′ (s, overlaps with H17a′), H13b′ (s)	
H13b′	H13a, H13a′, H13b		H13b, H13a′ (s), H13a (s)	C13, C13a, C13c
H17	H17a, H17a′ H18, H18a,a′	H17a, H18	H18a,a′ (s), H18 (s), H17a′ (s), H17a (s)	
H17a	H17a′	H17a′	H17a′ (s)	
H17a′	H17a	H17a	H17a (s)	
H17b	H17a, H17a′ H18a,a′	H17b′	H17a (s), H17a′ (s), H17b′ (s) H18 (s), H18a,a′ (s)	C17, C17a
H17b′	H17a, H17a′, H17b H18, H18a,a′		H17b, H17b (s), H18a,a′ (m) H17a (s), H17a′ (s)	C17, C17a
H18	H17a	H18a,a′, H19	H17a (s, overlaps with H3a′)	C17, C18b
H18a,a′	H17a, H17a′	H18	H17a (m), H17a′ (s, overlaps with H13a′)	C18, C18b, C19
H19	H17a, H17a′, H17b H17b′, H17, H18 H18a,a′, H20,20′	H18, H20,20′	H20,20′ (s), H17 (s) H18a,a′ (s), H18 (m)	
H20,20′	H17, H18, H18a,a′	H19	H17 (s), H18a,a′ (s), H18 (s), Me2 (s)	C1, C19, C18

[a] Primes refer to the downfield signal for resolved geminal protons pairs. s, m, and w refer to strong, medium, and weak NOESY cross peak intensities, respectively.

increment; 2.8-s repetition delay; 60-ms MLEV-17 spin-lock mixing period, preceded and followed by 2.5-ms trim pulses; 7.9-kHz spin-lock field, corresponding to 32-μs 90° plus widths; 6-Hz Gaussian and 90° shifted squared sine bell filtering in the t_2 and t_1 dimensions, respectively. Rotating frame Overhauser effect spectroscopy,[32-34] **ROESY**: 2 × 256 × 1024 raw data matrix size; 64 scans per t_1 increment; 2.4-s repetition delay; 85-ms continuous wave spin-lock period; 6.25-kHz spin-lock field strength, corresponding to 40-μs 90° pulse widths; 6-Hz Gaussian and 90°shifted squared sine bell filtering in the t_2 and t_1 domains, respectively. Nuclear Overhauser effect,[35,36] **NOESY**: 2 × 256 × 1024 raw data matrix size; 64 scans per t_1 increment; 2.8-s repetition delay period; 300-ms mixing period; 6-Hz Gaussian and 90° shifted squared sine bell filtering in the t_2 and t_1 domains, respectively. ^1H–^{13}C heteronuclear multiple quantum coherence,[37-42] **HMQC**: 2 × 218 × 1024 raw data matrix size; 128 scans per t_1 increment; 1.2-s repetition delay period; 128 scans per t_1 increment; 1.2-s repetition delay period; 33-μs 90° ^{13}C pulse widths; broad band Waltz-16 ^{13}C decoupling during the acquisition period; 3.5-ms defocusing/refocusing delay periods; 800-ms "Weft" period; 6-Hz Gaussian and 90° shifted squared sine bell filtering in the t_2 and t_1 dimensions, respectively. ^1H–^{13}C heteronuclear multiple-bond correlation,[43] **HMBC**: 1 × 200 × 1024 data matrix size; 512 scans per t_1 increment; 2.0-s repetition delay; 33-μs 90° ^{13}C pulse widths; 3.5-ms delay period for suppression of one-bond correlation signals; 40-ms delay periods for long-range couplings; 15° shifted squared sine bell filtering

(32) Bothner-By, A.; Stephens, R. L.; Lee, J. T.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811.
(33) Davis, D. G.; Bax, A. *J. Magn. Reson.* **1985**, *63*, 207.
(34) Bax, A.; Sklenar, V.; Summers, M. F. *J. Magn. Reson.* **1986**, *70*, 327.
(35) Jeneer, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *79*, 4546.
(36) Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 1980.
(37) Muller, L. *J. Am. Chem. Soc.* **1979**, *101*, 4481.
(38) Bax, A.; Griffey, R. G.; Hawkins, B. L. *J. Am. Chem. Soc.* **1983**, *105*, 7188.
(39) Bax, A.; Griffey, R. G.; Hawkins, B. L. *J. Magn. Reson.* **1983**, *55*, 301.
(40) Live, D. H.; Davis, D. G.; Agosta, W. C.; Cowburn, D. *J. Am. Chem. Soc.* **1984**, *106*, 6104.
(41) Bendall, M. R.; Pegg, D. T.; Doddrell, D. M. *J. Magn. Reson.* **1983**, *52*, 81.
(42) Bax, A.; Subramanian S. *J. Magn. Reson.* **1986**, *67*, 565.
(43) Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2093.

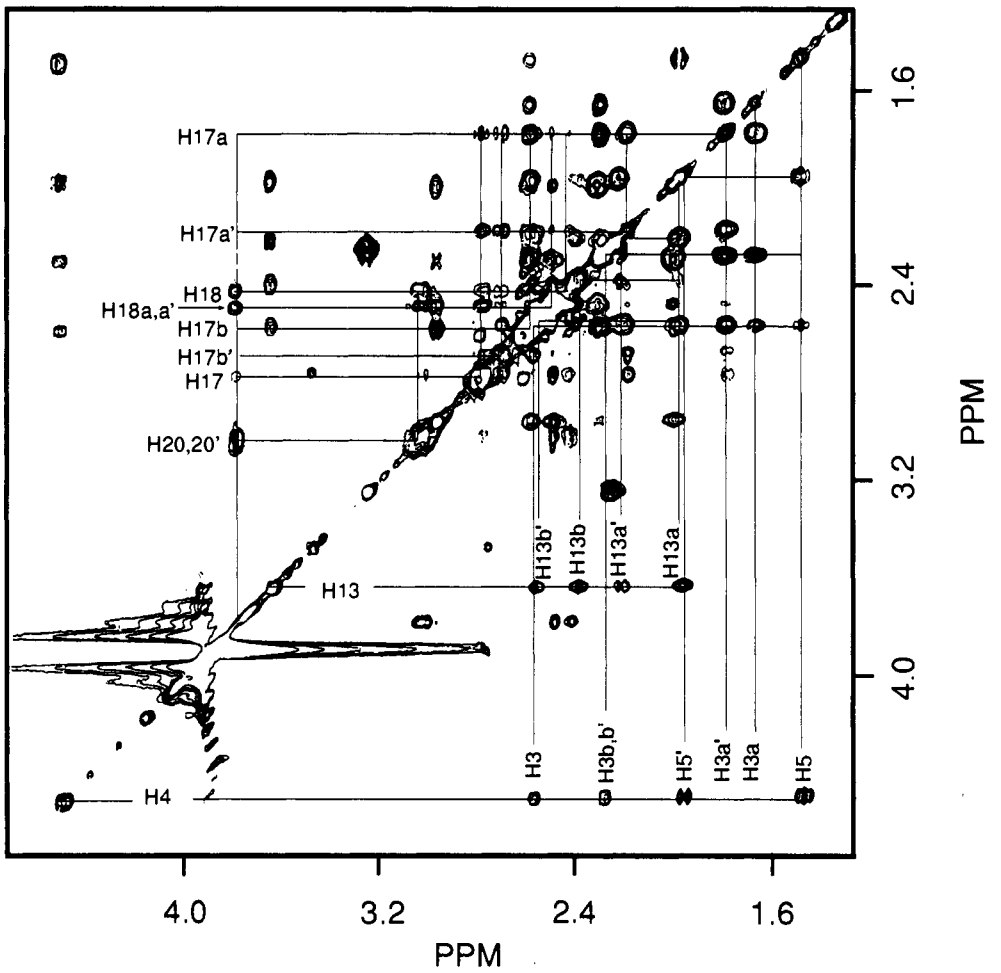


Figure 1. Portion of the 2D HOHAHA spectrum obtained for the 12,13-diepimeric form of coenzyme F430. Connectivities for *J*-networks associated with the H4, H19, and H13 protons are labeled.

in the t_2 dimension, and no digital filtering in the t_1 dimension.

NMR Methods and Spectra

^1H and ^{13}C NMR signal assignments were made following the procedures developed for coenzyme B_{12}[44,45] and more recently applied to the native form of coenzyme F430.[25] ^1H NMR signals were initially grouped into networks of scalar coupled protons (or *J*-networks), using HOHAHA (Figure 1) and 2QF-COSY (Figure 2) spectroscopies. Direct scalar connectivities were obtained from the COSY spectrum.

One-bond proton–carbon connectivities ($^1J_{CH}$) were determined with ^1H-detected, "BIRD-modified" heteronuclear multiple quantum coherence (HMQC) spectroscopy. The HMQC spectrum was useful for the identification of geminal protons since both ^1H signals correlate with a single carbon frequency, Figure 3. Long-range ($^2J_{CH}$, $^3J_{CH}$) ^1H–^{13}C connectivities were made with heteronuclear multiple-bond correlation (HMBC) spectroscopy. HMBC spectroscopy not only confirms the identification of long-range ^1H–^{13}C couplings through heteroatoms, but also enables the determination of primary structural features (Figure 4). The multiple-bond ^1H–^{13}C correlations observed for 12,13-diepimeric F430 are summarized in Table I. Through-space proton–proton dipolar connectivities are identified with two types of experiments: 2D nuclear Overhauser effect (NOESY, Figure 5) and 2D rotating frame Overhauser effect (ROESY) spectroscopies. The ROESY data are useful for distinguishing direct cross relaxation from spin diffusion effects; however, complexities in the ROESY peaks due to nonlinear spin-lock fields and Hartmann–Hahn effects make quantitation of dipolar relaxation rates more difficult to achieve. For quantitative measurement of dipolar relaxation, 2D-NOESY data were collected with mixing times of 10, 50, 100, 300, and 500 ms.

Signal Assignments

^1H NMR signals were assigned on the basis of direct and relayed scalar connectivities. Three independent "*J*-networks" (network of scalar-coupled protons) labeled in the HOHAHA spectrum result from scalar couplings associated with the H4, H19, and H13 protons (Figure 1). For example, in the HOHAHA spectrum the H4 proton signal at 3.63 ppm gives rise to cross peaks at 2.35 (H3), 2.55 (H3b,b'), 1.96 (H5'), and 1.49 (H5) ppm. These signals, plus additional signals observed at lower contour levels (H3a, 1.65 ppm; H3a', 1.79 ppm; see supplementary Figure S2), comprise a single *J*-network associated with the H4 proton. Direct scalar connectivities were identified in the DQF-COSY spectrum in Figure 2. All *J*-networks involving non-exchangeable protons were grouped and assigned in the same manner (see Table I).

Signals for carbons directly bonded to non-exchangeable protons were assigned from a ^1H–^{13}C HMQC spectrum (Figure 3). ^1H–^{13}C HMQC spectroscopy is particularly useful since signals for geminal protons are readily identified. For example, the well-separated H5 and H5' proton signals correlate with a single carbon frequency (C5) at 35.5 ppm. HMQC spectra also provided a ^1H–^{13}C correlation for severely overlapping geminal protons (H20,20', 3.00 ppm) at carbon frequency (29.1 ppm). At this stage, some signals had yet to be unambiguously assigned. For example, the ^1H and ^{13}C NMR signals of C2a and C7a position could not be distinguished since the isolated geminal protons do not provide further scalar coupling. However, a set of geminal

(44) Summers, M. F.; Marzilli, L. G.; Bax, A. *J. Am. Chem. Soc.* **1986**, *108*, 4285.
(45) Bax, A.; Summers, M. F.; Marzilli, L. G. *J. Am. Chem. Soc.* **1987**, *109*, 566.

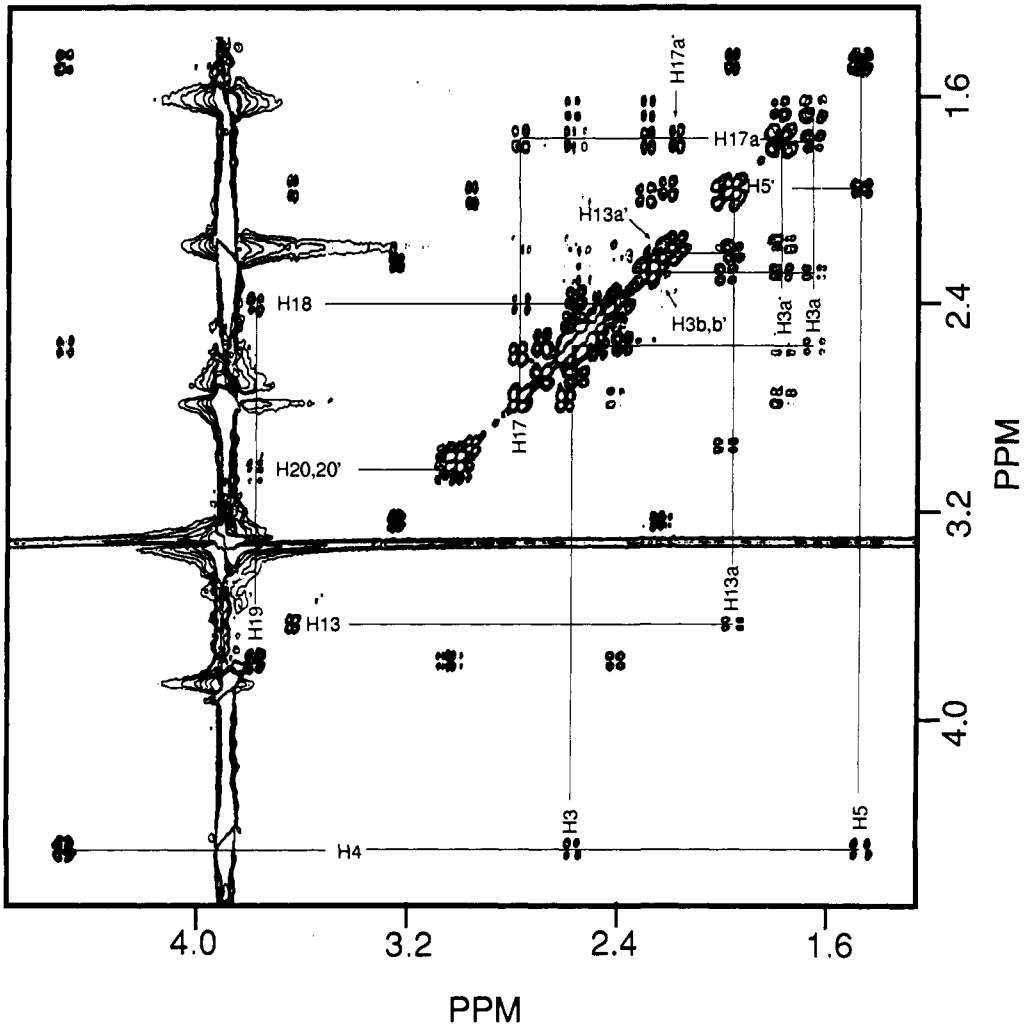


Figure 2. Portion of the 2D 2QF-COSY spectrum obtained for the 12,13-diepimeric F430 with positive and negative contour levels drawn. Direct scalar connectivities are labeled for the same *J*-networks shown in Figure 1.

protons (H7a,a') exhibits a strong NOE interaction with Me7, whereas the other set of geminal protons (H2a,a') exhibit strong dipolar cross relaxation with the Me2 protons. The spectral comparison with ^1H–^1H correlated spectroscopies (DQF-COSY and HOHAHA) and ^1H–^{13}C HMBC spectra (below) enabled to complete ^{13}C NMR signal assignments were made for all carbons with attached non-exchangeable protons (see Table II).

Most of the remaining ^{13}C signals were assigned from ^1H–^{13}C HMBC NMR data. A portion of the HMBC spectrum obtained for 12,13-diepimeric F430 is shown in Figure 4, and expansions are given in supplementary materials. Two- and three-bond ^1H–^{13}C long-range scalar connectivities were obtained from HMBC spectra. For example, HMBC correlation signals were observed from the H10 proton to four carbon atoms with ^{13}C chemical shifts of 57.5, 177.5, 169.1, and 47.8 ppm (Figure 4). The carbon signals at 57.5 ppm and 47.8 ppm were assigned previously to C8 and C12, respectively. The remaining signals at 177.5 ppm and 169.1 ppm are assigned to C9 and C12, respectively, based on HMBC connectivities involving H8 and H12. Because of the poor sensitivity resulting from relatively rapid proton T2 relaxation and weak multiple-bond CH scalar coupling, four ^{13}C signals representing carbon atoms C14, C15, C16, and C17c could not be assigned specifically, and tentative assignments for these carbons are included in Table II. ^1H and ^{13}C NMR signal assignments and spin connectivities are summarized in Tables II and III, respectively. In general, chemical shifts were in good agreement with shifts determined previously for native F430.

Table II. ^{13}C NMR Signal Assignments for Diepi-F430[a]

signal	diepi-F430	native F430[b]	signal	diepi-F430	native F430[b]
C17c	199.5	199.6	C17	51.9	51.8
C1	188.3	190.3	C7	51.7	51.5
C14	178.7[d]	180.5	C12	47.8	47.6
C3c	178.6[c]	180.2	C18	47.6	46.3
C18b	178.4[c]	179.7	C3	44.4	45.4
C8c	178.4[c]	179.7	C7a	44.3	44.8
C13c	178.2[c]	179.5	C2a	42.8	43.3
C9	177.5[c]	179.3	C12a	40.4	41.7
C12b	177.3[c]	177.8	C17b	39.1	39.3
C7b	177.3[c]	177.8	C18a	37.6	38.2
C2b	175.0	175.6	C5	37.0	37.5
C16	173.1[d]	174.8	C8b	33.9	35.2
C11	169.1	171.8	C3b	34.2	35.1
C15	109.5[d]	110.4	C13b	32.7	33.0
C10	99.9	100.3	C20	30.1	31.1
C6	93.9	94.1	C13a	29.6	29.5
C4	66.5	66.9	C17a	25.7	26.6
C19	65.7	65.2	C8a	23.3	24.0
C8	57.5	58.5	C3a	21.4	22.1
C2	55.3	55.9	C-Me2	20.2	20.6
C13	53.1	51.8	C-Me7	15.0	15.5

[a] Referenced to internal TFE-d_3, $\delta(^{13}C)$ = 61.5 ppm, $\delta(^1H)$ = 3.88 ppm. [b] From ref 25. [c] Signal assignments are tentative due to lack of adequate chemical shift dispersion in the HMBC spectra. [d] Signals not observed directly from HMBC spectra were tentatively assigned on the basis of ^{13}C chemical shift comparison.

Structure Determination

A. Stereochemical Assignments. Absolute stereochemical assignments for the A, B, and C rings were made previously on the basis of chemical and spectroscopic analyses,[7] and NOE intensities observed here are consistent with the earlier assignments. Relative stereochemical assignments of *R–R–S* (ORIGINAL) or *S–S–R* (REVERSE) were made for

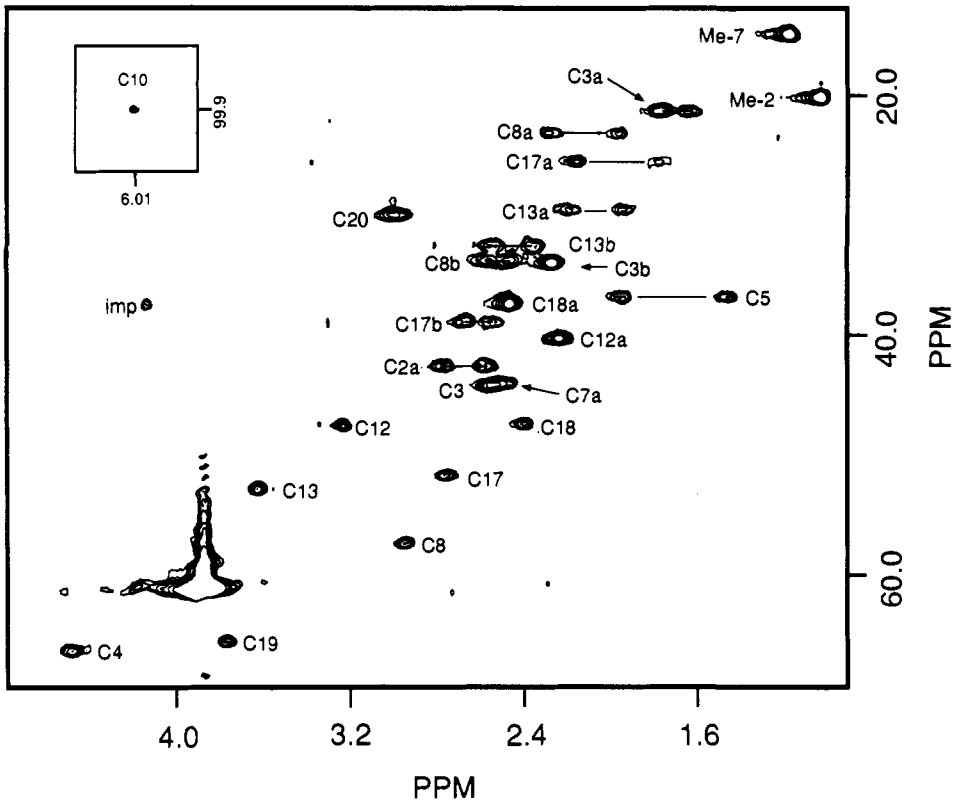


Figure 3. Portion of the 2D ^1H–^{13}C HMQC spectrum of 12,13-diepimeric F430. Horizontal lines connect signals of geminal protons. The H10–C10 correlation signal is included in the inset; imp refers to an impurity.

Table III. ^1H NMR Signal Assignments for 12,13-Diepimeric F430[a] and Native F430[b]

assignment	diepimer	native	assignment	diepimer	native
Me2	1.05	1.08	H10	6.01	5.91
Me7	1.19	1.18	H12	3.23	3.13
H2a	2.59	2.60	H12a,a′	2.24	2.67
H2a′	2.77	2.73	H13	3.63	3.91
H3	2.56	2.60	H13a	1.96	1.71
H3a	1.65	1.65	H13a′	2.21	1.93
H3a′	1.79	1.77	H13b	2.35	2.09
H3b,b′	2.28	2.26	H13b′	2.55	2.29
H4	4.48	4.45	H17	2.76	2.83
H5	1.47	1.48	H17a	1.79	1.78
H5′	1.95	1.95	H17a′	2.17	2.19
H7a,a′	2.48	2.49	H17b	2.55	2.58
H8	2.95	2.86	H17b′	2.71	2.63
H8a	1.97	1.98	H18	2.40	2.38
H8a′	2.30	2.26	H18a,a′	2.47	2.45
H8b	2.48	2.44	H19	3.77	3.55
H8b′	2.57	2.56	H20,20′	3.00	3.02

[a] Chemical shifts are referenced to internal TFE-d_3 (ppm). [b] From ref 25.

the C17–C18–C19 carbons, respectively, based on relative NOE intensities for ring protons.[46] It was impossible to determine the absolute stereochemistries of these groups from our NOESY data due to partial overlap and strong coupling observed for the C20 geminal protons. An initial attempt to determine the stereochemistry of C19 was made using 2D NOESY back-calculations[46] for models generated with both stereochemical configurations. Calculations for the REVERSE model indicated that this structure could be made consistent with the experimental data, but that internal distortions (reflected by covalent penalties) were obtained. On the other hand, the ORIGINAL structures could be manipulated to be consistent with data without inducing internal strain. These findings indicated that the assignments were originally by Eschenmoser and co-workers[8] were correct. However, the X-ray structural and NMR analyses of 12,13-diepimeric F430M demonstrate conclusively that the original stereochemical assignment is incorrect.[11] We believe that our inability to identify the correct stereoisomer was due to improper use of "floating chirality". Thus, stereochemical assignments for geminal methylene protons on proteins are routinely made by allowing the geminal protons to swap positions (or "float") during structure refinement. However, in our F430 studies, we allowed sterically non-equivalent groups (H-19 and the C18 moiety) to float. Inherent in this approach was the assumption that the correct stereochemistry would be the one with the lowest penalty (minimum van der Waals and bond-angle violations). Clearly, this is not the case, and the use of floating chiralities for NMR-based stereochemical assignment of sterically non-equivalent groups should be avoided. In addition, it is apparent that parameters for primary restraints, including bond angles and van der Waals radii, that are routinely employed for protein structure determination are not appropriate for the modelling of small molecules. In particular, modifications must be made to account for van der Waals overlap for atoms in non-aromatic, six-membered rings and in fused ring systems (see below).

The absolute stereochemistries for the A-, B-, and C-rings were assigned based on earlier chemical and NMR studies of native and 12,13-diepi-F430M,[8] and the absolute stereochemistry of the C17–C18–C19 group was assigned based on recent NMR and X-ray studies of 12,13-diepi-F430M.[11] These stereochemical assignments were used for the structural modelling described below.

B. Distance Restraints for Covalent Structure, Distance restraints to account for covalent bonding, electronic conjugation, and van der Waals radii were defined on the basis of published X-ray structural data for relevant molecules.[47,48] Covalent distances and angle restraints for atoms of the A-, B-, and C-pyrrole rings of the corphin macrocycle were obtained by averaging published values for related moieties of the equatorial macrocycle of coenzyme B_{12} and methyl B_{12}. Bond lengths used in the initial DG/SA structure generation were classified into several different bond types (mean ± standard deviation): C(sp3)—C(sp3), 1.56 ± 0.05 Å; C(sp2)=C(sp2), 1.35 ± 0.05 Å; C(sp3)—C(sp2), 1.53 ± 0.03 Å; C(sp2)=N(sp2), 1.329 ± 0.03 Å; C(sp3)—N(sp2), 1.434 ± 0.03 Å; and C(sp2)=O(sp2), 1.25 ± 0.02 Å (see Table IV).

Bond angle restraints for the corphin ring were also generated by averaging data published for B_{12} (mean ± standard deviation): N-(sp2)–C(sp2)–C(sp3), 113.31 ± 2.0°; C(sp2)–C(sp3)–C(sp3), 102.20 ±

(46) Olson, K. D.; Won, H.; Wolfe, R. S.; Hare, D. R.; Summers, M. F. *J. Am. Chem. Soc.* **1990**, *112*, 5884.

(47) Hodgkin, D. C.; Pickworth, J.; Robertson, J. H.; Prosen, R. J.; Sparks, R. A.; Trueblood, K. N. *Proc. R. Soc. London* **1959**, *A251*, 306.

(48) Rossi, M.; Glusker, J. P.; Randaccio, L.; Summers, M. F.; Toscano, P. J.; Marzilli, L. M. *J. Am. Chem. Soc.* **1985**, *107*, 1729.

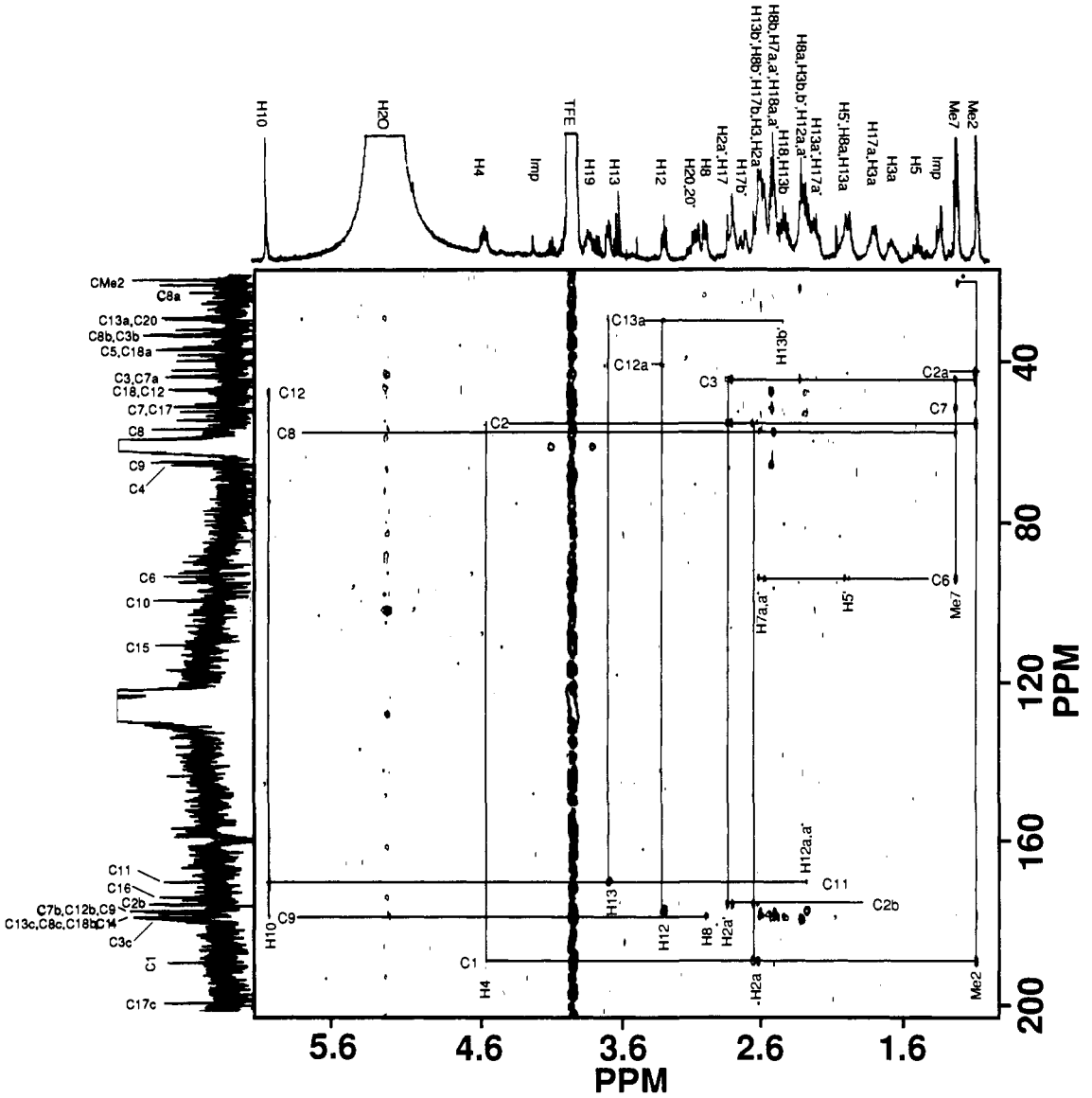


Figure 4. Portion of the 2D HMBC spectrum of 12,13-diepimeric F430 showing two- and three-bond scalar connectivities used to assign carbon signals and identify covalent connectivities. Signals labeled * and imp are due to the incompletely suppressed zero quantum (one-bond) coherences and sample impurities, respectively. Expanded regions of the HMBC spectra are labeled in supplementary Figures S3 and S4.

1.5°; C(sp3)–C(sp3)–C(sp3), 106.19 ± 1.5°, and C(sp3)–N(sp2)–C(sp3), 113.44° ± 2.0°; C(sp2)–C(H)–C(sp2), 126 ± 2°.

Interatomic covalent distances and bond angles for atoms of the fused lactam ring were obtained from X-ray structural data published for a vitamin B_{12} degradation product that contained a similar fused lactam moiety.[47,48] As experimentally observed in coenzyme B_{12} for the exocyclic bond angles associated with C2, C3, and C4 carbon atoms, a bond angle (116.5 ± 2.5°) was used while Csp3–Csp3–H (or H–Csp3–H) was set to ideal tetrahedral geometry as 109.5°. For the C(sp2) carbon of the five-membered ring, a bond angle (123.33 ± 1.5°) for two exocyclic bond angles, N(sp2)–C(sp2)–C(sp3) and C(sp3)–C(sp2)–C(sp3), was applied. Covalent distances for the fused six-membered ring were obtained from X-ray crystallographic results published for steroid and sugar molecules with related eletronic structure.[49,50]

Covalent distances for side-chain groups were also derived from averaged X-ray diffraction data for B_{12} as follows: C(sp3)–C(sp3), 1.56 ± 0.05 Å; C(sp3)–O(sp2), 1.35 ± 0.03 Å; C(sp3)–N(sp2), 1.28 ± 0.04 Å. Standard C–H, N–H, and O–H were employed: C(sp3)–H, 1.09 Å; C(sp2)–H, 1.08 Å; N(sp2)–H, 1.00 Å; O–H, 1.09 Å. All covalent distances employed for structure determination are summarized in Table IV. Values in this table obtained by averaging X-ray data for B_{12}-related structures are reported as the mean, and standard deviations for each bond type are included in the footnotes of Table IV.

In addition to the above covalent bond restraints, additional restraints were included to enforce planarity for electronically conjugated regions of the macrocycle, including the N2, C9, C10, C11, N3, C14, C15, C16, N14, and C17c, and for atoms of the exocyclic lactam moiety. Amide and carboxyl groups were restrained to minimize van der Waals repulsion with methylene protons on the neighboring carbons.

In several cases, atom pairs were allowed to approach to distances smaller than the sums of their van der Waals radii. This is an important point for researchers that intend to use software designed for studies of proteins in studies of small molecules. For example, standard radii for carbons of the six-membered ring necessarily led to a flattening of the ring due to the van der Waals repulsion terms. As such, carbon pairs of the six-membered ring separated by three bonds were allowed to approach to 2.85 Å (determined as the minimum separation observed in X-ray structures of several sugars and steroids),[49,50] which enabled the maximum sugar pucker observed from X-ray crystallographic studies of relevant sugar and steroid structures. All planarity restraints and restraints used to allow sub-van der Waals approach of atoms are summarized in Table IV.

Ni–N bond distances of 1.85–1.95 Å are typically observed for low-spin, square-planar Ni(II) complexes.[18,19] As such, two separate sets of structures were generated using Ni–N distances of 1.85 Å and 1.95 Å.

C. Interproton Distance Restraints. Interproton distance restraints were generated on the basis of cross peak intensities in the 2D NOESY

(49) The Chemical Society, London, In Tables of Interatomic Distances and Configuration in Molecules and Ions, and Interatomic Distances Supplement; Burlington House: W.1, London, 1965 (printed by Eyre and Spottiswoode Ltd., The Thanet Press).

(50) Wheatley, P. J. Physical Methods in Heterocyclic Chemistry, Academic Press: New York, 1972; pp 479–520.

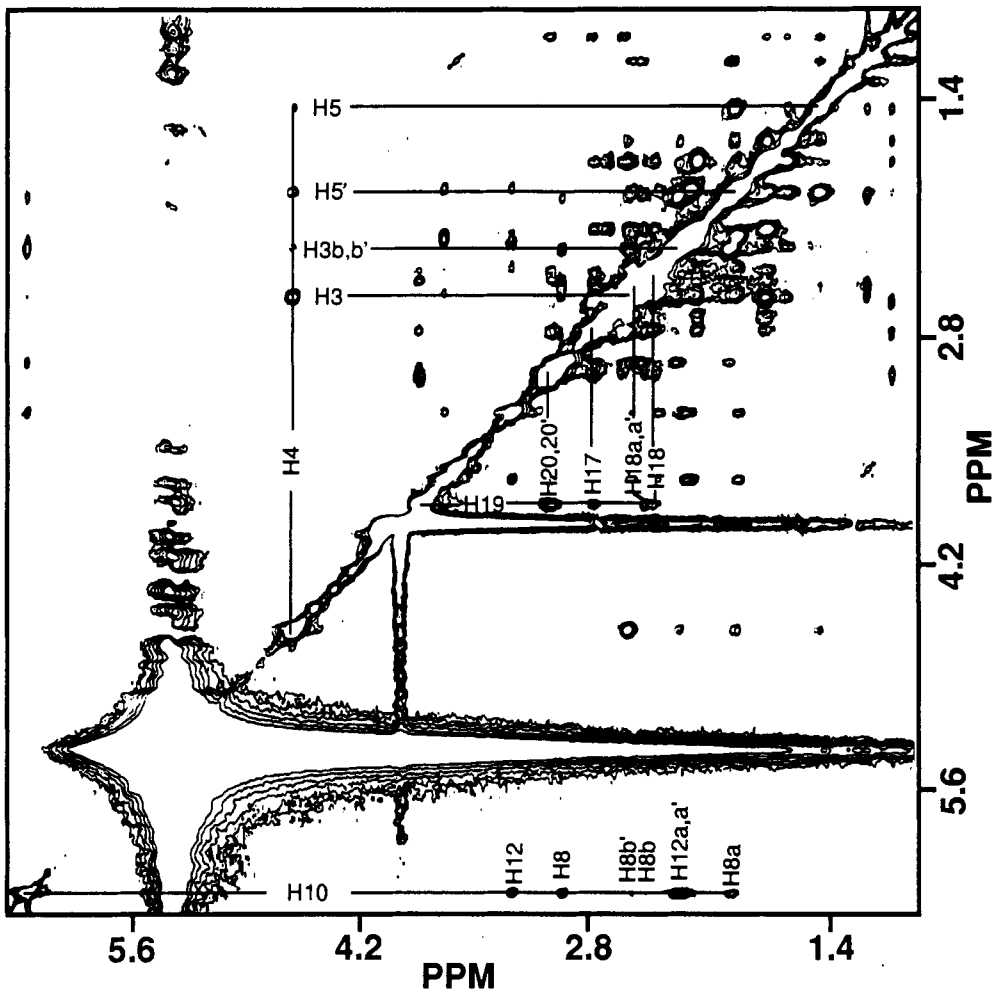


Figure 5. 2D NOESY spectrum (τ_m = 300 ms) obtained for 12,13-diepimeric F430 in deuterated trifluoroethanol (TFE) solution. Cross peaks associated with the H10, H4, and H19 corphin ring protons that were important for determining conformational elements are labeled.

Table IV. Atomic Parameters and Primary Restraints Used for Distance Geometry Calculations

bonded atom distance restraints[a]		planarity restraints[b]	
bond type/atom	distance/radii, Å	atoms	distance, Å
C–H (sp^3–s)	1.09	Ni–C1	2.75–2.90
C–H (sp^2–s)	1.08	Ni–C4	2.84–2.98
N–H (sp^2–s)	1.00	Ni–C6	2.84–2.98
C–N (sp^3–sp^3)	1.39	Ni–C9	2.75–2.90
C–N (sp^2–sp^2)	1.33	Ni–C11	2.81–2.95
C–C (sp^3–sp^3)	1.56	Ni–C14	2.81–2.95
C–C (sp^2–sp^2)	1.43	Ni–C16	2.75–2.90
C–C (sp^3–sp^2)	1.53	Ni–C19	2.84–2.98
C–O (sp^2–sp^2)	1.25	C15–C17a	2.85–3.91
Ni–N (dsp^2–sp^2)	1.85	C16–C17b	2.85–3.92
		C17–C17c	2.85–3.90
H	0.95	C14–Oc17c	2.60–2.70
N	1.30	C13–Oc17c	2.30–2.50
C	1.50	C13–Oc17c	2.70–3.10
O	1.30	H13–Oc17c	1.50–infinite
Ni	0.70	C13a–Oc17c	1.50–infinite

[a] Values for atomic radii and for C–H and N–H distances are from refs 22 and 23. Other interatomic distances are reported as averages from X-ray crystallographic results for coenzyme B_{12} and B_{12} related compounds (refs 47 and 48), with standard deviations as follows (Å): C–N (sp^3–sp^3), ±0.02; C–N (sp^2–sp^2), ±0.03; C–C (sp^3–sp^3), ±0.05; C–C (sp^2–sp^2), ±0.03; C–C (sp^3–sp^2), ±0.03; C–O (sp^2–sp^2), ±0.01. [b] Restraints to enforce planarity of conjugated atoms were calculated using simple trigonometric identities.

spectra following the general procedure developed for structural studies of proteins.[20,21] Distance restraints of 2.0–2.5 Å, 2.0–3.5 Å, and 2.0–4.5 Å were used to represent strong, medium, and weak cross peak intensities observed in a NOESY spectrum obtained with a 300-ms mixing period. Cross peaks resulting from spin diffusion were differentiated from direct dipolar cross peaks based on analysis of 2D ROESY spectra obtained with MLEV-17 mixing periods ranging from 50 to 150 ms, and on the basis of NOE data collected with mixing times of 10, 50, 100, 300, and 500 ms.

Stereochemical assignment of several prochiral geminal protons were made on the basis of relative proton NOE intensities. For example, the well-separated signals due to the geminal C5 protons both exhibited strong NOE cross relaxation with the Me7 protons. In addition, H4 exhibited direct NOE cross relaxation with the downfield C5 proton signal, and exhibited spin diffusion with the upfield C5 proton. The upfield and downfield C5 proton signals could therefore be assigned to pro-*S* and pro-*R* protons, respectively. Other geminal protons for which prochiral assignments could be made readily include those at the C8a, C13a, C13b, C17a, C2a, and C3a. All other geminal proton signals exhibited severe overlap, precluding prochiral signal assignment. The complete set of interproton distance restraints are summarized in Table V.

D. Structure Generation and Refinement. Two sets of 20 structures each were generated by standard hybrid distance geometry/simulated annealing calculations with the program DSPACE. These calculations refine against a penalty function (penalty = sum of the square of the deviations between the input and calculated distances). More sophisticated energy-minimization-type calculations were not performed. In the first set (DG(primary)), restraints to account for primary geometry and electronic conjugation were included, but no interproton distance restraints based on NOESY data were included. These calculations were carried out in order to assess the contribution of primary structural restraints on the degree of DG/SA structure convergence and resolution. A second set of DG/SA structures (DG(+NOE)) was generated which included NOE-derived distance restraints. For both sets of structures, the Ni–N distances were fixed at 1.85 Å. For the DG(+NOE) calculations, the primary and experimental distance restraints were given equal weighting.

DG/SA structures were generated and refined using a modification of an approach described previously.[23] Distance pairs were embedded randomly into three-space and subjected to simulated annealing (SA) and

Table V. NOE Restraints Used for the Generation of Distance Geometry Structure DG21-60(+NOE)[a]

H10 restraints		H17 restraints	
H10–H12	2.0–3.5	H17–H17a	2.0–3.5
H10–H8	2.0–3.5	H18 restraints	
H10–H8a′	2.0–3.5	H18–H17a	2.0–3.5
H4 restraints		H18–H20	2.0–3.5
H4–H3	2.0–2.5	H19 restraints	
H4–H3b	2.0–4.5	H19–H20	2.0–2.5
H4–H5′	2.0–2.5	H19–H17	2.0–3.5
H8 restraints		Me2 restraints	
H8–H8a′	2.0–2.5	Me2–H20	2.0–3.5
H8–H8b	2.0–3.5	Me2–H2a	2.0–3.5
H8–H8b′	2.0–3.5	Me2–H2a′	2.0–3.5
H12 restraints		Me2–H3a	2.0–3.5
H12–H13a	2.0–3.5	Me2–H3a′	2.0–3.5
H12–H13b	2.0–3.5	Me2–H5	2.0–3.5
H12–H13	2.0–3.5	Me7 restraints	
H13 restraints		Me7–H7a	2.0–3.5
H13–H13a	2.0–2.5	Me7–H5	2.0–2.5
H13–H13a′	2.0–2.5	Me7–H5′	2.0–3.5
H13–H13b	2.0–4.5	Me7–H8a	2.0–2.5
H13–H13b′	2.0–4.5	Me7–H8b	2.0–3.5

[a] Restraints are in unit of Å.

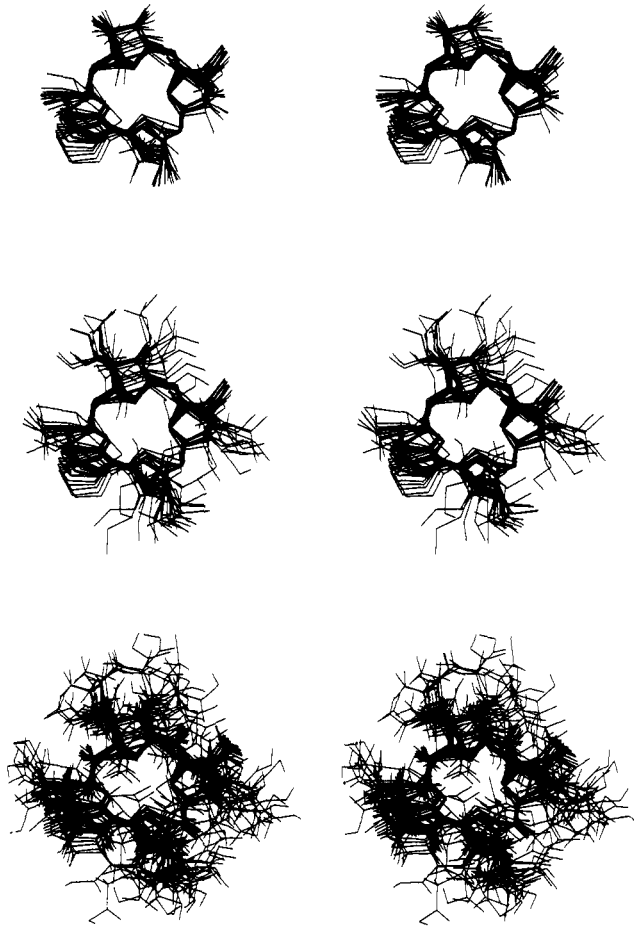


Figure 6. Superposition of 20 DG/SA structures generated with only primary distance restraints. Superpositions were obtained for the macrocyclic ring atoms only (top), all atoms except hydrogens (middle), and all atoms (bottom).

conjugated gradient minimization (CGM). After minimization to a moderate target penalty (ca. 0.3 Å^2), the initial DG/SA structure was saved, and new DG/SA structures were generated by performing two 10-Å randomizations of atom positions, followed by SA and CGM refinement. Initial DG(primary) and DG(+NOE) structures exhibited penalty values in the range 0.04–0.3 Å^2. Further refinement was achieved by application of variable velocity simulated annealing (to

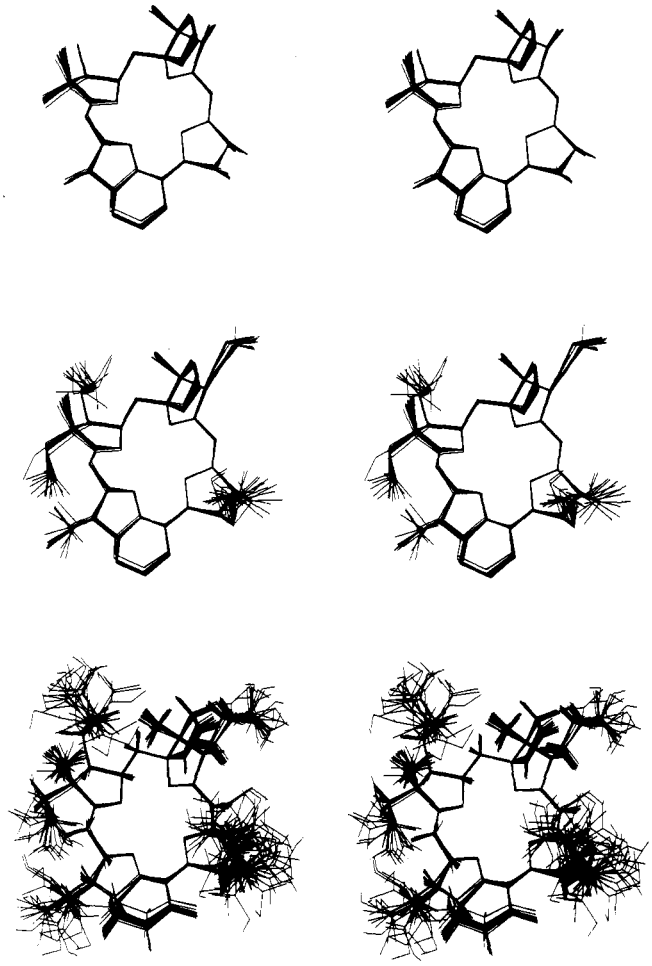


Figure 7. Superposition of 20 DG/SA structures generated with primary and NOE-derived "loose" distance restraints (see text). For comparison with Figure 6, superpositions were performed for the macrocyclic ring atoms only (top), all atoms except hydrogens (middle), and all atoms (bottom).

maximum penalty values of 10–20 Å^2), SHAKE (to penalties of ca. 10 Å^2), and CGM algorithms.[23] Final DG(primary) and DG(+NOE) structures exhibited penalty values in the range 0.036–0.05 Å^2 and 0.048–0.065 Å^2, respectively.

Results and Discussion

Degree of Convergence. Forty final DG/SA structures generated with primary restraints only (DG1-20(primary)) and with primary and experimental NOE restraints (DG21-40(+NOE)) were evaluated in terms of their degree of structural convergence. The degree of convergence provides an assessment of the ability of the input parameters to define a single conformation. Convergence was assessed in terms of the pairwise root-mean-square deviations (RMSD) obtained upon superposition of macrocyclic atoms: N1, C5, N2, C10, N3, C15, N4, C20. For the DG(primary) structures, pairwise RMSD values in the range of 0.08 to 1.17 Å were obtained, whereas superposition of all atoms afforded pairwise RMSD values in the range of 1.19 to 3.42 Å. As shown in Figure 6, the 20 DG(primary) models do not converge to a unique conformation. Although the structures exhibit ruffled corphin ring conformations, the conformations of the partially saturated pyrrole units (rings A–D) vary significantly among the 20 DG/SA structures, as reflected by the relatively high pairwise RMSD values.

On the other hand, good conformational convergence was observed for the macrocyclic ring atoms of the DG(+NOE) structures, with pairwise RMSD values in the range 0.02 to 0.15 Å, Figure 7. As discussed below, the five-membered pyrrole units also converged to a unique conformation. The side-chain groups associated with C3a, C12a, C13a, and C18a did not converge,

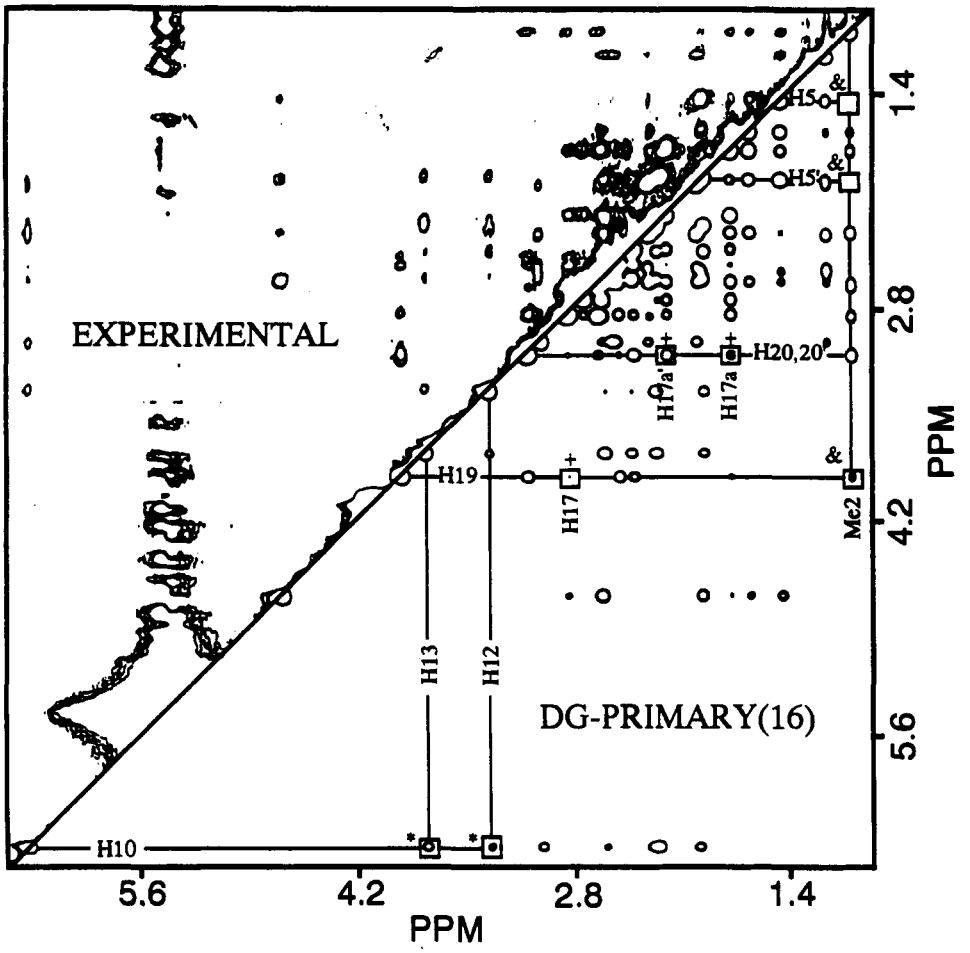


Figure 8. Comparison of the experimental 2D NOESY spectrum of 12,13-diepi-F430 (above diagonal) with the back-calculated 2D NOE spectrum of DG-16(primary), a representative DG/SA structure generated with primary restraints but no NOE-derived interproton distance restraints (below diagonal). Several of the obvious discrepancies between the spectra are labeled (*, +, and & symbols refer to discrepancies associated with the H10, H19 and H20,20′, and Me2 protons, respectively), and reflect improper puckering of the five-membered rings of the corphin macrocycle (see text).

as expected since no stereospecific NOE restraints were included for protons on these groups, Figure 7. Pairwise RMSDs obtained upon superposition of all atoms were in the range 0.67 to 1.29 Å.

As described in detail below, the DG(+NOE) structures exhibit a highly saddle-shaped ring conformation. This conformation was predicted for 12,13-diepi-F430 based on X-ray structural studies of Ni(II)–hydrocorphin model compounds.[14,18]

Evaluations of DG/SA Structures Based on 2D NOESY Back-Calculations. To determine how well the final DG/SA structures represent the experimental NOESY data, 2D NOESY back-calculations were performed.[22-24] The algorithm (BKCALC) employs what is basically an emprical two-parameter fitting of the spectral density, one parameter (k_{cr} = 75 s^{-1}; determined by fitting the NOE buildup of geminal protons) which governs the cross relaxation rate, and the other (k_{zl} = 1 s^{-1}, except for methyl protons where k_{zl} = 3 s^{-1} was employed; determined by measuring the loss of integrated, total spectral magnetization as a function of increasing mixing time) which governs the "z-leakage" rate. Z-leakage accounts for loss of z-magnetization due to all processes other than cross relaxation (e.g., multiple quantum, chemical exchange, etc.). As described previously, this approach accurately accounts for spin diffusion.[22-24]

The back-calculated 2D NOESY spectra (300 ms mixing time) for representative DG/SA structures generated without and with NOESY-derived distance restraints are compared with the experimental NOESY spectrum in Figures 8 and 9, respectively. Figure 8 contains the back-calculated spectrum obtained for DG-16(primary), a structure that exhibits maximum RMS deviation when compared to the DG(+NOE) structures. Ignoring signals for the β side-chain protons, several discrepancies are observed. Consider the A-ring first, where moderate-to-strong cross peaks were observed for the Me2-H5,H5′ proton pairs in the experimental NOESY spectrum. Similar cross peaks are lacking in the back-calculated spectrum, Figure 8. In addition, the back-calculated spectrum contained a spin diffusion cross peak of moderate intensity for the Me2–H19 protons, whereas a similar peak was not observed experimentally. The spin diffusion peak resulted from the fact that Me2 and H19 are very close (<3.97 Å) to an H20 proton in the DG-16(primary) model. The combined back-calculated data reflect the fact that the A-ring of DG-16(primary) exists in an Me2(equatorial)–H3(equatorial)–H4(axial) conformation. The fact that the back-calculated NOESY data for these protons do not match the experimental NOESY data provides good evidence that the A-ring conformation of DG-16(primary) does not accurately reflect the actual, solution-state conformation of the A-ring of 12,13-diepi-F430.

Similar discrepancies were observed for protons of the C- and D-rings of DG-16(primary). In the back-calculated spectrum, a strong cross peak is observed for the H10–H13 proton pair and a weak peak is observed for the H10–H12 pair. This peak pattern reflects the fact that the C-ring of DG-16(primary) exists in an H12(axial)–H13(axial) conformation. Experimentally, a strong H10–H12 cross peak is observed, and no H10–H13 signal is observed, indicating that the C-ring conformation of DG-16-(primary) is incorrect. Finally, DG-16(primary) exhibited a strong cross peak for the H20,20′–H17a,a′ protons and a very weak cross peak for the H17–H19 proton pairs, reflecting the fact that the D-ring exists in a H17(equatorial)–H18(equatorial)–H19(equatorial) conformation. Strong NOE cross relaxation was experimentally observed for the H17–H19 protons, and no cross relaxation was observed for the H20,20′–H7a,a′ protons, again demonstrating that the D-ring conformation of DG-16(primary) is not correct.

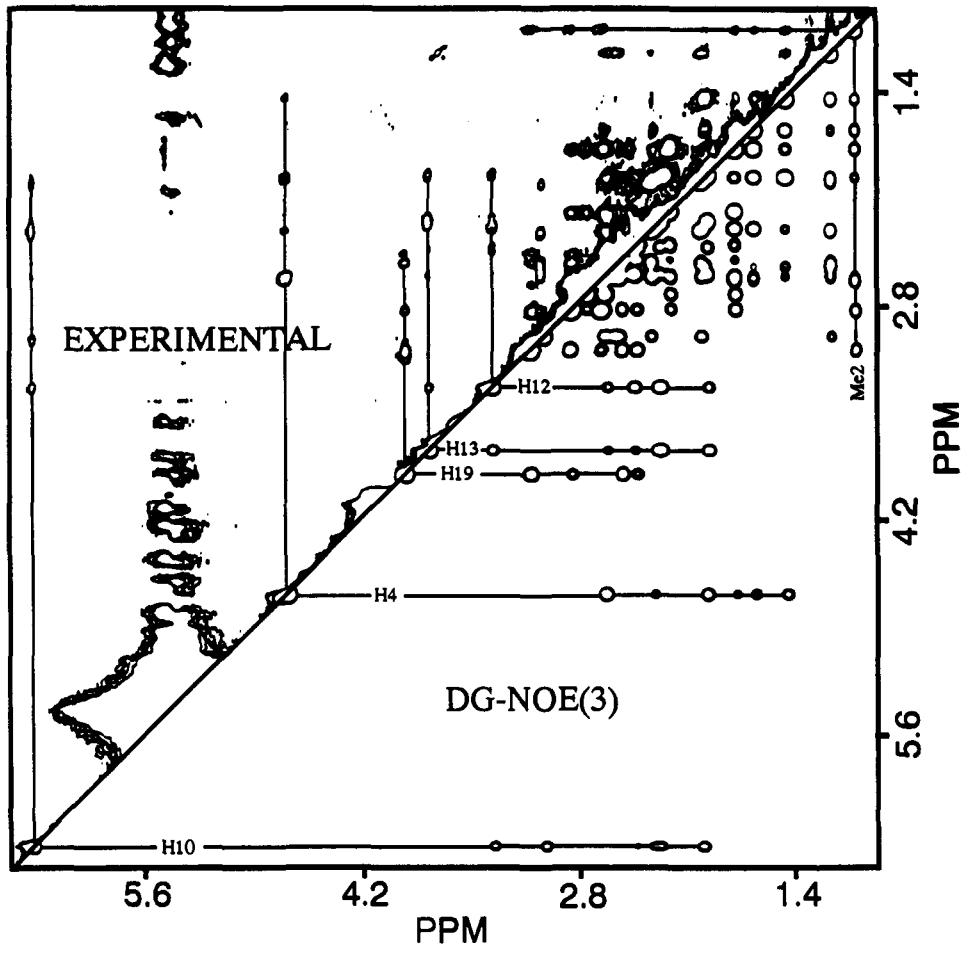
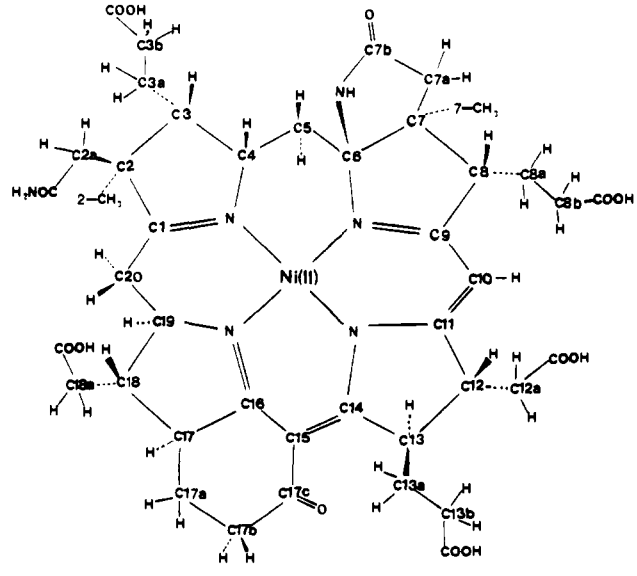


Figure 9. Comparison of the experimental 2D NOESY spectrum of 12,13-diepi-F430 (above diagonal) with the back-calculated 2D NOE spectrum of DG-3(+NOE), a representative DG/SA structure generated with primary and "loose" NOE-derived distance restraints (below diagonal). The match between the experimental and back-calculated spectra is quite good, demonstrating that the DG-3(+NOE) structure more accurately represents the experimental NMR data compared with the DG-16(primary) structure; see Figure 8 for comparison.

As indicated in Figure 6, the DG(primary) structures exhibit a wide range of ring conformations involving many combinations of different ring pucker conformations. DG-16(primary) represents the "worst case" structure, exhibiting maximum RMS deviation when compared to the DG(+NOE) structures. As shown in Figure 7, the DG(+NOE) structures exhibit good convergence, and except for signals of a few side-chain groups, back-calculated NOE spectra for the DG(+NOE) structures are internally self-consistent. The back-calculated spectrum obtained for DG-3(+NOE), the structure with lowest total penalty, is compared with the experimental NOESY spectrum in Figure 9. In contrast to the back-calculated spectra of the DG(primary) structures, cross peak patterns and intensities in the back-calculated spectra of the DG(+NOE) structures closely match those of the experimental NOESY spectrum. Even the NOE build-up profiles as a function of increasing mixing period give good fits to the experimental data (see supplementary figures). The conformations of the A-, C-, and D-rings are Me2(axial)–H3(axial)–H4 (equatorial), H12 (equatorial)–H13(equatorial), and H17(axial)–H18(axial)–H19(axial), respectively, for all of the 20 DG(+NOE) models.

Influence of the Ni–N Covalent Distance on the DG(+NOE) Conformation. Twenty additional DG(+NOE) structures were generated in a manner identical with that described above for DG(21–40), except that the Ni–N bond distance restraint was set to 1.95 Å. As indicated above, low-spin, square-planar Ni(II) complexes typically exhibit Ni–N distances in the range 1.85 to 1.95 Å, and these calculations allowed us to determine the influence of the Ni–N distance parameter on DG/SA conformation. The conformations of these new structures were essentially identical with the conformations of DG(21–40), with pairwise RMSDs for 40 superpositioned structures (20 with 1.85-Å Ni–N distance restraints and 20 with 1.95-Å Ni–N distance restraints)

Chart I



in the range 0.019 to 0.15 Å (see supplementary Figure S10). It appears that the NOE restraints lock the macrocycle into a unique conformation, and that lengthening the Ni–N distances results in a displacement of the Ni above the 4N plane but does not lead to a significant change in the NOE-restrained conformation of the macrocycle.

Comparison with the X-ray Structure of 12,13-Diepi-F430M. The 20 DG/SA structures generated with the inclusion of

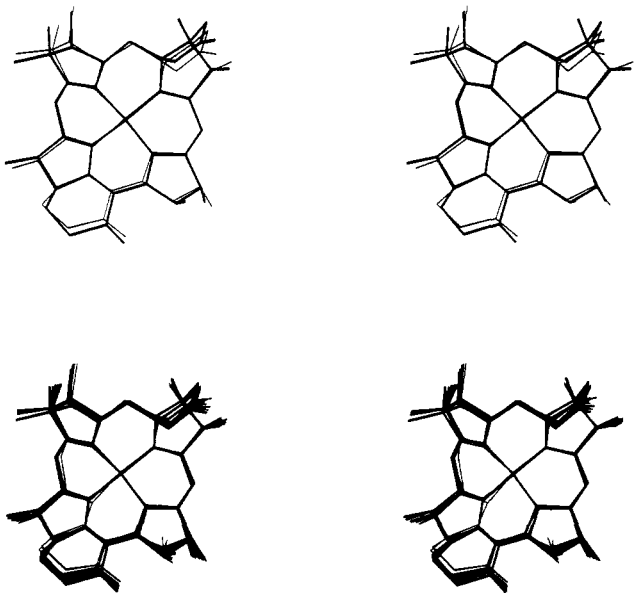


Figure 10. Superposition (matching of the 24 corphin ring atoms) of the 20 DG(+NOE) models (thin lines) with the crystal structure of 12,13-diepi-F430 pentamethyl ester (heavy line) (bottom), and superposition of the DG(+NOE)(3) model (thin line) with the crystal structure of 12,13-diepi-F430 pentamethyl ester (fat line) (top).

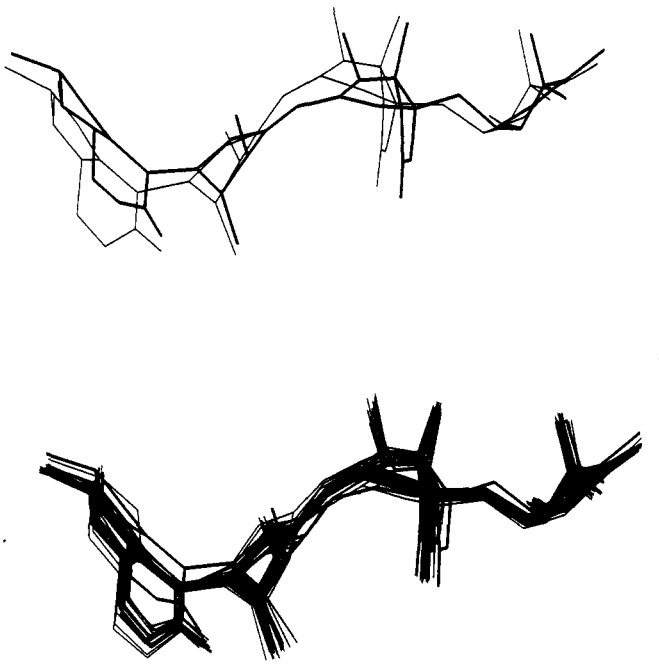


Figure 11. Cylinder projection (radius 2.5 Å) of the superposition (matching of the 24 corphin ring atoms) of the 20 DG(+NOE) models (thin lines) with the crystal structure of 12,13-diepi-F430 pentamethyl ester (fat line) (bottom), and cylinder projection of the superposition of the DG(+NOE)(3) model (thin line) with the crystal structure of 12,13-diepi-F430 pentamethyl ester (fat line) (top).

NOE-derived distance restraints DG(+NOE)(1–20) are assessed in terms of the conformational variability *among* the 20 DG(+NOE) structures and are compared with the recently determined crystal structure of 12,13-diepi-F430M. As a measure of the conformational similarity between two structures (derived from NMR or X-ray data), the RMS distance $\langle d \rangle$ between equivalent atoms is used. The quantity $\langle d \rangle$ is computed (for a specified set of atoms) after translationally and orientationally matching the two structures subject to the condition that $\langle d \rangle$ (for the same or a different set of atoms) is a minimum.

Corphin Ring Atoms. Superpositions of the 20 DG(+NOE) structures with the X-ray structure after matching the 24 corphin ring atoms (N1 to N4, C1 to C20) are shown in Figure 10 (bottom). The agreement between X-ray conformation and DG(+NOE) conformations ranges between $\langle d \rangle$ = 0.25 Å (DG-4(+NOE)) and 0.31 Å (DG-11(+NOE)), with an average for all DG(+NOE) structures of 0.28 Å. The agreement between pairs of DG(+NOE) structures shows a much larger spread in $\langle d \rangle$ values ranging between 0.03 Å and 0.36 Å, with an average agreement between any two DG(+NOE) structures of 0.145 Å. The full matrix of $\langle d \rangle$ values for all pairs of DG(+NOE) structures and the crystal structure is included in supplementary Table S1. The corresponding data for the corphin atoms plus the (nonhydrogen) atoms considered to be conformationally restricted (i.e., all atoms directly attached to the corphin ring plus the atoms of two extra rings) are included in the supplementary material.

The combined data indicate that there is a small but systematic difference between the X-ray structure and any of the DG(+NOE) structures. In terms of $\langle d \rangle$ values, this systematic difference is about as large as the maximum conformational variation within the set of DG(+NOE) structures. In order to facilitate a structural interpretation of the difference between X-ray structure and NOE structures, we chose the conformation DG(+NOE)(3) as the "representative" conformation for the sample of NMR conformations: DG(+NOE)(3) is the conformation with the property that the sum of the $\langle d \rangle$ values with the 19 other DG(+NOE) conformations has the lowest value among all 20 conformations, and it is the conformation with the lowest total penalty (see above). A superposition between this "representative" NMR conformation and the X-ray conformation is shown in Figure 10 (top).

It has been demonstrated in the past that a suitable representation for the nonplanarity of a porphinoid ring is a *cylinder projection*. Cylinder projections of the superpositions of the X-ray structure with 20 DG(+NOE) structures and with the "representative" NMR model, DG(+NOE)(3), are shown in Figure 11. The wave-shaped appearance of this representation of the corphin ring is the result of a saddle-shaped deformation, induced by a contraction of the corphin ring to accommodate the stereochemical requirement of the "small" central Ni(II) atom.[18] Evidently, the X-ray conformation shows a smaller *amplitude* of the trace of the projection, indicative of a larger degree of saddle deformation in the DG(+NOE) conformations. The d_m values, which quantifies the extent of saddle deformation, reflects this situation: d_m is 0.785 Å for the X-ray structure and 0.965 Å for DG(+NOE)(3). The average d_m for all DG(+NOE) structures is 0.928 Å, with a minimum of 0.843 Å and maximum 0.981 Å.

The larger saddle deformation is associated with a more contracted structure, as manifested by shorter transannular N–N distances for the NMR conformations. DG(+NOE)(3): $d_{N1\cdot N3}$ = 3.55 Å, $d_{N2\cdot N4}$ = 3.70 Å; X-ray: $d_{N1\cdot N3}$ = 3.69 Å, $d_{N2\cdot N4}$ = 3.69 Å. Yet, the Ni–N distances are very similar: the average of the four Ni–N distances is 1.856 Å for DG(+NOE)(3) and 1.858 Å for the X-ray structures. This is due to the fact that the Ni(II)–N distances were fixed at 1.85 Å in the DG(+NOE) structures. No coordination geometry restraints were included for Ni in the calculations, and the DG(+NOE) structures exhibit a much less planar Ni–N coordination geometry compared with the X-ray structures; perpendicular deviations from a plane through the four pyrrolic N-atoms in the sequence Ni, N1, N2, N3, N4 are the following. X-ray: –0.018, 0.204, –0.194, 0.203, –0.212; DG(+NOE)(3): –0.182, 0.331, –0.321, 0.350, –0.361.

Side-Chain Atoms. As described above, the side-chain groups of the DG structures did not converge to unique conformations. One could reasonably argue that the side chains undergo rapid conformational fluctuations on the NMR time scale. However, the only accurate statement that can be made is that the scatter observed for the side chains in the DG/SA structures (Figure 7) results from a lack of quantitative NOE information. This lack of information is due largely to signal overlap observed for the geminal-β protons, which precluded stereochemical assignments. The degeneracy of these β-methylene signals *might* be attributable to rapid thermal motion on the NMR time scale. As expected, the DG(+NOE) structures compared less favorably with the X-ray structures when all atoms were included in the fits (see supple-

mentary Table S1 and Figure S10).

Summary

An approach has been described for structural modelling of relatively small molecules by NMR. The first step involves setting up primary restraints to account for the covalent structure of the molecule of interest. For 12,13-diepi-F430, covalent angles and distances were derived from X-ray crystallographic data published for related coenzyme B_{12} derivatives. In general, standard van der Waals radii were employed, with the exception that atoms of the five- and six-membered rings of the macrocycle were allowed to approach to distances less than the sums of their van der Waals radii in order to maximize conformational sampling. It is essential to utilize proper primary restraint parameters when extending algorithms designed for structural modelling of biopolymers to small molecules, and one must demonstrate that the primary restraints alone allow for sufficient conformational flexibility. Structures generated by our distance geometry approach using only primary restraints exhibited significant conformatiomal diversity, with each of the unsaturated five-membered pyrrole units and the six-membered ring existing in several low-penalty conformations. Back-calculated NOESY spectra for these DG/SA structures were shown to deviate significantly compared to the experimental NOESY spectra.

Structures generated subsequently with the addition of "loose" NOE-derived distance restraints converged to a unique conformation, and the similarity of the back-calculated and experimental NOESY spectra indicates that the DG(+NOE) structures are reasonable representatives of the solution-state conformation of 12,13-diepi-F430. The DG(+NOE) structures compare favorably with the X-ray structure of 12,13-diepi-F430M, with pairwise RMSDs for the 24 macrocyclic ring atoms of ca. 0.28 Å. The DG(+NOE) and X-ray structures exhibit a similar corphin ring conformation that is characterized by an extreme saddle-shaped ring deformation, which was predicted on the basis of X-ray studies of Ni(II)-containing hydrocorphinoid complexes.[16-19]

The NMR-based approach has several limitations; in particular, regardless of the quality of NMR data, the "resolution" of the NMR structures will certainly never approach the high resolution obtainable by X-ray diffraction techniques. In addition, the NMR approach relies intimately on knowledge of the primary structure, including atomic stereochemistry and electronic conjugation, and without full knowledge of these parameters the implementation of the NMR-based approach should be avoided. Nevertheless, for molecules that have proven difficult or impossible to crystallize, such as native F430, the NMR approach provides an attractive alternative for structural modelling. Applications to other small-molecule systems for which high-resolution X-ray structural results are available will establish the reliability and ultimate potential of this technique.

Acknowledgment. Financial support from the Petroleum Research Fund, administered by the American Chemical Society (to M.F.S.), NSF Grant DMB 86-13679 and NIH Grant AI12277 (to R.W.), and technical support from Bernie Duffy (UMBC), Victor Gabriel (UIUC), and Scott Smith (UIUC) are gratefully acknowledged. C.K. acknowledges support from the Austrian Science Foundation.

Registry No. 12,13-Diepimeric F430, 142186-78-5.

Supplementary Material Available: Figures containing UV/vis spectra of F430 and 12,13-diepimeric F430 in H_2O, expansions of a 2D HOHAHA spectrum with the labeled *J*-network associated with the H4 proton at lower contour levels, portions of HMBC spectra, experimental and back-calculated NOE build-up profiles, and superposition of DG21-60(+NOE) structures as well as tables of pairwise RMSD values for DG/SA and X-ray structures (14 pages). Ordering information in given on any current masthead page.

Electron Transfer and Bond Breaking. Examples of Passage from a Sequential to a Concerted Mechanism in the Electrochemical Reductive Cleavage of Arylmethyl Halides

Claude P. Andrieux, Annie Le Gorande, and Jean-Michel Savéant*

Contribution from the Laboratoire d'Electrochimie Moléculaire de l'Université de Paris 7, Unité Associée au CNRS N° 438, 2 place Jussieu, 75251 Paris Cedex 05, France.
Received January 29, 1992. Revised Manuscript Received April 1, 1992

Abstract: A systematic investigation of the kinetics of the electrochemical reduction of a series of 11 arylmethyl halides in acetonitrile and *N*,*N*'-dimethylformamide reveals a striking change in the reductive cleavage mechanism as a function of the energy of the π^* orbital liable to accept the incoming electron. With ring-substituted nitrobenzyl chlorides and bromides, a stepwise mechanism involving the intermediacy of the anion radical takes place. When slightly less electron-withdrawing substituents, such as nitrile or ester groups, are involved, the reaction occurs via a concerted electron transfer–bond breaking mechanism. This is also observed with the unsubstituted benzyl chloride and bromide as well as with 9-anthracenylmethyl chloride. The main factor then governing the thermodynamics and kinetics of the reductive cleavage is then the dissociation energy of the bond being broken. There is an excellent agreement between the predictions of the recently developed model of dissociative electron transfer and the experimental data both in terms of the quadratic character of the activation–driving force relationship and the magnitude of the intrinsic barrier. In this connection, procedures utilizing the difference in cyclic voltammetric peak potential between two compounds bearing the same nucleofugal group to estimate the difference in their bond dissociation energies are discussed.

Electron transfer to molecules is very often accompanied by the breaking of an existing bond and/or the formation of a new bond. In the first case, where a radical and an anionic fragment are formed upon electron transfer to a neutral molecule, an im-

0002-7863/92/1514-6892$03.00/0 © 1992 American Chemical Society