tion based on NMR data. *J Computer Aided Mol Design* **1**: 219–241, 1987.
9. Lautz J, Kessler H, Blaney JM, Scheek RM and van Gunsteren WF, Calculating three-dimensional molecular structure from atom-atom distance information: Cyclosporin A. *Int J Pep Protein Res* **33**: 281–288, 1989.
10. Lautz J, Kessler H, van Gunsteren WF, Weber HP and Wenger RM, On the dependence of molecular conformation on the type of solvent environment: A molecular dynamics study of cyclosporin A. *Biopolymers*, in press.
11. Neuhaus D and Williamson M, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. VCH, New York, 1989.
12. James TL, Borgias BA, Bianucci AM, Jamin N, Pattabiraman N, Suzuki E, Zhou N and Zon G, Molecular structure determination via complete relaxation matrix analysis of 2D NOE spectra and molecular mechanics: Application to DNA fragments. In: *NMR Spectroscopy in Drug Research* (Eds. Jaroszewski JW, Schaumburg K and Kofod H), pp. 358–375. Munksgaard, Copenhagen, 1988.
13. Boelens R, Koning TMG and Kaptein R, Determination of biomolecular structure from proton-proton NOE's using a relaxation matrix approach. *J Mol Struct* **173**: 299–311, 1988.
14. Boelens R, Koning TMG, van der Marel GA, van Boom JH and Kaptein R, Iterative procedure for structure determination from proton-proton NOEs using a full relaxation matrix approach. Application to a DNA octamer. *J Magn Reson* **82**: 290–308, 1989.
15. Gehrke M, Dissertation, Frankfurt am Main, 1989.
16. van Gunsteren WF and Berendsen HJC, *Groningen Molecular Simulation (GROMOS) Library Manual*, pp. 1–229. Biomos B.V., Nijenborgh 16, NL 9747 AG Groningen, 1987.
17. Kessler H, Griesinger C and Wagner K, Peptide conformations. 42. Conformation of side chains in peptides using heteronuclear coupling constants obtained by two-dimensional NMR spectroscopy. *J Am Chem Soc* **109**: 6927–6933, 1987.
18. Kessler H, Bats JW, Griesinger C, Koll S, Will M and Wagner K, Peptide conformations. 46. Conformational analysis of a superpotent cytoprotective cyclic somatostatin analogue. *J Am Chem Soc* **110**: 1033–1049, 1988.
19. Seebach D, Struktur und Reaktivität von Lithiumenolaten, von Pinakolon zur selektiven C-Alkylierung von Peptiden—Schwierigkeiten und Möglichkeiten durch komplexe Strukturen. *Angew Chem* **100**: 1685–1715, 1988.
20. Seebach D, Thaler A and Beck AK, Solubilization of peptides in non-polar organic solvents by the addition of inorganic salts: Facts and implications. *Helv Chim Acta* **72**: 857–867, 1989.
21. Kessler H, Gehrke M and Griesinger C, Two-dimensional NMR spectroscopy: Background and overview of the experiments. *Angew Chem Int Ed Engl* **27**: 490–536, 1988.
22. Kessler H, Gemmecker G, Köck M, Osowski R and Schmieder P, Improvement of carbon detected H,H-TOCSY experiments by employing a DEPT transfer. *Magn Reson Chem* **28**: 62–67, 1990.
23. Bax A, Davis DG and Sarkar SK, An improved method for two-dimensional heteronuclear relayed-coherence transfer NMR spectroscopy. *J Magn Reson* **63**: 230–234, 1985.
24. Kessler H, Kerssebaum R, Klein AG, Obermeier R and Will M, Synthese und Konformationsanalyse von cyclischen Alanin-Analogen des Thymopoietins durch NMR-Spektroskopie und Molecular-Dynamics-Rechnungen im Vakuum und in Lösung. *Liebigs Ann Chem* 269–294, 1989.
25. Szilagyi L and Jardetzky O, α-Proton chemical shifts and secondary structure in proteins. *J Magn Reson* **83**: 441–449, 1989.

Biochemical Pharmacology, Vol. 40, No. 1, pp. 173–175, 1990.
Printed in Great Britain.

0006-2952/90 $3.00 + 0.00

Structures from NMR distance constraints

Although advances in NMR have elevated structural determination of peptides and small proteins in solution to the level of a routine tool [1], the nature of the structural information obtained is not often appreciated. The best known method for structural analysis, X-ray crystallography, yields results in Cartesian coordinate space—the average position of each heavy atom in a molecule. In contrast, NMR studies reveal information in distance and torsional space. Distances between pairs of atoms are determined from nuclear Overhauser effects (NOEs) [2], paramagnetic enhancement of relaxation [3], or hydrogen bonding. Torsion angle constraints are derived from coupling constants and Karplus-type relationships [4]. Like X-ray, these reflect the ensemble of conformations present during the time frame of the measurement [5]. Unlike X-ray, however, these observables can be measured even in regions of structure that are highly disordered and have no characteristic conformation in solution. Generation of atomic coordinates from this information relies on several computational methods, most commonly distance geometry [6] and molecular dynamics [7, 8]. Although these methods have been shown to be robust in the elucidation of solution conformation at low resolution, it is clear that any effort to correlate biological activity with solution conformation and dynamic properties is vitally dependent on the discovery of all possible structures consistent with the NMR data and on some means of filtering quality information (regions which do have a characteristic conformation) from data compromised by conformational heterogeneity.

The issue of completeness in conformational sampling remains an open one. Although distance geometry has been suggested to randomly sample conformational space in much the same fashion as Monte Carlo [9], some investigators have suggested that in the absence of sufficient constraints, it produces fully extended structures that are nearly identical [10]. This suggests that conformational sampling by distance geometry is biased and that care must be exercised in overinterpreting the results of computations for structures or regions of structures that are underdetermined. The complexity of molecular dynamics calculations means that they are time consuming, and consequently it is not possible to run simulations long enough to ensure that all of conformational space is sampled. In addition, weighting of the NMR-derived distance constraints sufficient to ensure convergence to a folded structure may predispose against complete sampling. The problem of becoming trapped in a local minimum has limited the use of optimization methods in generating

structures, and is a legitimate concern in molecular dynamics. Dynamical simulated annealing [11] involving high temperature simulations, has been used to increase sampling by enhancing the likelihood of transit between minima. Molecular dynamics [7] (when used alone), Monte Carlo [9], and minimization in torsional space [12] have been shown to suffer from path dependency in the implementation of NMR-derived constraints. All were forced to resolve the issue by defining local structure initially and subsequently enforcing constraints having the greatest impact in determining the overall protein fold.

Aside from sampling issues, most computational methods implicitly assume homogeneity of information. Distance geometry and molecular dynamics, by forcing a compromise between the potential function describing the energy of the system and the experimental constraints, can produce structures with distortions in bond angles and bond lengths which result from distributing local errors over the entire ensemble. Approaches such as DISMAN [12] and Monte Carlo which assume ideal geometry and deal with torsional variables circumvent a part of this problem, but use of global target minimization procedures still allows contamination of precisely determined substructures with data from under determined regions of the structure.

Problems in conformational sampling and data homogenization can best be approached by a systematic evaluation of all possible torsional variables. The very nature of this approach means that the result is path independent. Being deterministic, it will identify *all* conformations at a given scan increment consistent with the data set. Small regions can be analyzed individually, and inconsistencies in experimental data, which are a clear hallmark of conformer averaging, will be determined automatically. After fixing the range of acceptable values for each torsional variable, and determining acceptable combinations, one can arrive at a final structure by building larger and larger aggregates having few or no internal degrees of freedom. Each unique path through variable space can be walked without duplication in a systematic fashion. Our application of systematic search to NMR data for cyclosporin-A (CSA) suggests that this approach does offer advantages over currently existing computational methods, particularly as the number and quality of available constraints increase.

Methods and Results

Solution conformations of CSA, an eleven-residue immunosuppressive cyclic peptide, were determined from an NMR data set in dimethyl sulfoxide (DMSO) [13]. The crystal structure of iodocyclosporin [14] was used to generate a model for the backbone of CSA, and the SEARCH module of the SYBYL software package [15] was used to generate conformers by scanning ϕ and ψ angles in 10° increments. Conformers were screened for unfavorable van der Waals' contacts using radii previously calibrated for proteins and peptides [16]. To speed the calculation, the CSA model was analyzed initially as several small subproblems. Two- or three-residue fragments were evaluated to determine ϕ, ψ angle ranges consistent with fifty-five short-range NOEs and vicinal coupling constants. Ultimately, the molecule was analyzed in two major fragments sharing common amide bonds. The loop fragment consisted of Ala7, D-Ala8, MeLeu9, MeLeu10, and MeVal11, and the sheet fragment included MeBmt1, Abu2, Sar3, MeLeu4, MeVal5, and MeLeu6. Torsional ranges for ϕ and ψ determined in small subsearches, five long-range NOEs, and four hydrogen bonds were used to constrain the search of these major fragments. Distances linking atoms at the ends of the loop fragment were monitored for every valid conformation. The resulting distance map described all possible relative orientations between terminal atoms in the loop fragment. Since the loop and sheet fragments shared common atoms, this map could be used as a constraint in the subsequent analysis of the sheet fragment. In a final calculation, the distance map generated from the sheet fragment was used to constrain a search of the loop fragment. Use of a distance map from one fragment to constrain the search of the other fragment ensured that not only the NMR-derived constraints were met, but also that the ring closure requirement was satisfied. The sheet and loop fragments were then combined to generate complete CSA structures by matching conformer pairs that possessed identical distance map points.

Two and one-half million conformations were found to be available to the eleven CSA backbone ϕ, ψ pairs. Two distinct conformational families were found to be available to Sar3. One family was a βII′ turn ($\phi 3 = 60°$, $\psi 3 = 120°$) and constituted ~94% of the valid conformations. This is equivalent to the βII′ structures reported for CSA [13] and its iodo derivative [14] in the crystalline state. The remaining 6% corresponded to a βI turn ($\phi 3 = -50°$, $\psi 3 = -60°$). Both families were found to be energetically acceptable. Residues other than Sar3 were conformationally less variable, having similar ϕ,ψ values regardless of the Sar3 conformation.

The results of this study contrast with two recent structural studies based on the same NMR data set. A 30-psec molecular dynamics simulation of CSA [17] at 300°K found no βI structure, and the authors concluded that the backbone solution conformation of CSA was similar to that of the crystal. Both the molecular dynamics simulation and the search calculations unexpectedly suggest that the sheet region of CSA is more flexible than the loop. A second study using distance geometry [18] alone and in combination with molecular dynamics also failed to find the alternative conformation.

Discussion

Our studies of the solution conformation of cyclosporin suggest that commonly employed computational methods fail to identify all structures consistent with NMR observations. While the 2.5 million conformations found may seem a startlingly large number, if one assumes only two values for each torsional variable, a total of over 4 million combinations (2^{22}) is possible. Thus, individual variables in CSA were actually tightly constrained. Any application of systematic search to proteins will require many precisely known constraints—a limitation that is rapidly disappearing as efforts to rigorously interpret NMR data continue. Refinement of distance constraints by spectral simulation [19] and incorporation of realistic motional models, stereospecific assignments [20] and determination of heteronuclear coupling constants [21] may result in cases where structures are actually overdetermined. Even so, there will still be situations in which the data are insufficient to dictate convergence to a single structure. In those cases, the inherent inability of NMR to resolve an ensemble of conformations can be ameliorated by a knowledge of the conformational preferences of amino acids. Databases are undoubtedly better at defining local conformation in the absence of experimental data than algorithms currently in use.

Center for Molecular Design
Washington University
St. Louis, MO 63130-4899
U.S.A.

DENISE D. BEUSEN
HIROSHI IIJIMA
GARLAND R. MARSHALL*

* Correspondence: Prof. Garland Marshall, Center for Molecular Design, Washington University, Lopata Hall, Box 1099, One Brookings Drive, St. Louis, MO 63130-4899.

REFERENCES

1. Wüthrich K, Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science* **243**: 45–50, 1989.

2. Noggle JH and Schirmer RE, *The Nuclear Overhauser Effect.* Academic Press, New York, 1971.
3. Mildvan AS, NMR studies of the interaction of substrates with enzymes and their peptide fragments. *FASEB J* **3**: 1705–1714, 1989.
4. Karplus M, Vicinal proton coupling in nuclear magnetic resonance. *J Am Chem Soc* **85**: 2870–2871, 1963.
5. Jardetzky O, On the nature of molecular conformations inferred from high-resolution NMR. *Biochim Biophys Acta* **621**: 227–232, 1980.
6. Havel T and Wüthrich K, A distance geometry program for determining the structures of small proteins and other macromolecules from nuclear magnetic resonance measurements of intramolecular ^1H–^1H proximities in solution. *Bull Math Biol* **46**: 673–698, 1984.
7. Clore GM, Brunger AT, Karplus M and Gronenborn AM, Application of molecular dynamics with interproton distance restraints to three-dimensional protein structure determination. *J Mol Biol* **191**: 523–551, 1986.
8. Clore GM and Gronenborn AM, Determination of three-dimensional structures of proteins in solution by nuclear magnetic resonance spectroscopy. *Protein Engineering* **1**: 275–288, 1989.
9. Bassolino DA, Hirata F, Kitchen DB, Kominos D, Pardi A and Levy RM, Determination of protein structures in solution using NMR data and IMPACT. *Int J Supercomputer Appl* **2**: 41–61, 1988.
10. Metzler WJ, Hare DR and Pardi A, Limited sampling of conformational space by the distance geometry algorithm: Implications for structures generated from NMR data. *Biochemistry* **28**: 7045–7052, 1989.
11. Nilges M, Clore GM and Gronenborn AM, Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett* **229**: 317–324, 1988.
12. Braun W and Go N, Calculation of protein conformations by proton-proton distance constraints. *J Mol Biol* **186**: 611–626, 1985.
13. Loosli H-R, Kessler H, Oschkinat H, Weber H-P, Petcher TJ and Widmer A, Peptide Conformations. The conformation of cyclosporin A in the crystal and in solution. *Helv Chim Acta* **68**: 682–704, 1985.
14. Petcher TJ, Weber H-P and Ruegger A, Crystal and molecular structure of an iodo-derivative of the cyclic undecapeptide cyclosporin A. *Helv Chim Acta* **59**: 1480–1488, 1976.
15. SYBYL Theory Manual, Tripos Associates, Inc., 1699 South Hanley Drive, St. Louis, MO 63144, 1989.
16. Iijima H, Dunbar JB Jr and Marshall GR, Calibration of effective van der Waals atomic contact radii for proteins and peptides. *Proteins: Struct Funct Genet* **2**: 330–339, 1987.
17. Lautz J, Kessler H, Kaptein R and van Gunsteren WF, Molecular dynamics simulations of cyclosporin A: The crystal structure and dynamic modelling of a structure in apolar solution based on NMR data. *J Computer Aided Mol Design* **1**: 219–241, 1987.
18. Lautz J, Kessler H, Blaney JM, Scheek RM and van Gunsteren WF, Calculating three-dimensional molecular structure from atom-atom distance information: Cyclosporin A. *Int J Pept Protein Res* **33**: 281–288, 1989.
19. Montelione GT, Winkler ME, Rauenbuehler P and Wagner G, Accurate measurements of long-range heteronuclear coupling constants from homonuclear 2D NMR spectra of isotope-enriched proteins. *J Magn Reson* **82**: 198–204, 1989.
20. Guntert P, Braun W, Billeter M and Wuthrich K, Automated stereospecific ^1H NMR assignments and their impact on the precision of protein structure determinations in solution. *J Am Chem Soc* **111**: 3997–4004, 1989.
21. Summers MF, Hare D, South TL and Kim B, Structure of a retroviral zinc finger: 2D NMR spectroscopy and distance geometry calculations on a synthetic finger from HIV-1 nucleic acid binding protein, p7. *J Cell Biochem Suppl* **13A**: 17, 1989.