$$\boxed{\text{Ph-}t\text{Bu}} \rightleftharpoons \boxed{t\text{Bu-Ph}} \quad \Delta\Delta G^\circ \qquad (4)$$

In Figure 5 $\Delta\Delta G$'s calculated by using the equilibrium constants listed in Table II are plotted as a function of the concentration of the salt. The fact that $\Delta\Delta G^\circ$ is enhanced as a function of concentration of the salt indicates that the solvation structure also is accordingly changed. The detailed nature of this change is not quite clear from the numbers in Table III because the change of $\Delta\Delta G^\circ$ by the addition of the salt is the result of complex combination of the change of $\Delta\Delta H^\circ$ and $\Delta\Delta S^\circ$. The hydrophobicity parameter estimated on the basis of the partition of an organic material to water and octanol by Hansch and Leo[19,20] is 1.98 and 1.96 for *tert*-butyl and phenyl, respectively. Although the *tert*-butyl group seems to be slightly more hydrophobic, the difference seems not to be enough to account for the present experimental results. Hydrophobic interactions, which are effective in the salt effect, could be different from the quantitative measure obtained from the extractability by organic solvents especially in the case of the aromatic group. This seems to be in line with the fact that the aromatic group needs a characteristic parameter in molecular mechanics calculations of the structure of inclusion compounds.[21]

Experimental Section

ESR Spectroscopy. Probes **1** and **2** were prepared from phenyl lithium and α-2,4,6-(trimethoxyphenyl)-*N*-*tert*-butyl nitrone (**3**) and phenyl *N*-*tert*-butylnitrone (**4**) by the method described previously.[10] All solutions were prepared by using distilled water treated with Millipore Milli Q system. The pH of the solution was adjusted by using an appropriate buffer: pH 7–9, NaH_2PO_4, NaOH; pH 11–12, Na_2HPO_4, NaOH; pH 12–13, KCl, NaOH. All chemicals except **3**, **4**, α-, and γ-cyclodextrin were obtained from Aldrich Chemical Co. α- and γ-cyclodextrin were purchased from Sigma Chemical Co. Spin traps **3** and **4** are available in these laboratories. Sample solutions whose probe concentrations were less than 10^{-4} M were loaded in Pyrex tubes, 1 mm i.d. and 2 mm o.d. ESR spectra were obtained with a Bruker ER 200D-SRC spectrometer. The field modulation width was set less than 0.016 mT, and the incident microwave power was set less than 6.3 mW. The temperature was controlled by using a Bruker ER 4111T variable-temperature unit. Relative concentrations of each spectral species are determined by using computer spectrum simulation. ESR lines with nitrogen nuclear spin 0 and 1 (center- and high-field wing of nitrogen hfs) were used for the fitting.

Calculation of pK_a. In order to simplify the calculation, all hydroxyl groups of cyclodextrin are assumed to dissociate at a single pH; then the dissociation equilibrium present in this system is simplified as

$$CD \rightleftharpoons CD^- + H^+ \qquad K_a = [CD^-][H^+]/[CD] \qquad (5)$$

where CD, CD^-, and H^+ show cyclodextrin, cyclodextrin anion, and hydronium ion, respectively. Thus, the concentration of cyclodextrin in this equilibrium is[16]

$$C_1 = C_0/(1 + K_a/[H^+]) \qquad (6)$$

where C_0 denotes the initial concentration of cyclodextrin. From the inclusion equilibrium of the complex, the association constant of the complex is expressed as[10]

$$K_{CD} = (r_1/r_0)[C_1 - [R]r_1/(r_1 + r_0)]^{-1} \qquad (7)$$

where r_0 and r_1 denote the relative concentration of the free and induced species and [R] shows the initial probe concentration. Since $C_1 \gg [R]$, the term including [R] can be neglected. Combining eq 6 and eq 7, one can obtain the following equation:

$$pK_a = pH - \log (r_0 C_0 K_{CD}/r_1 - 1) \qquad (8)$$

Acknowledgment. Support of this research provided by the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. We also thank Dr. Coit M. DuBose for the availability of α-2,4,6-(trimethoxyphenyl)-*N*-*tert*-butyl nitrone.

(19) Hansch, C.; Leo, A. *Substituents Constants for Correlation Analysis in Chemistry and Biology*; Wiley: New York, 1979.
(20) For recent review of hydrophobic interaction, see: Jiang, X.-K. *Acc. Chem. Res.* **1988**, *21*, 362.
(21) For example, see: Menger, F. M.; Sherrod, M. J. *J. Am. Chem. Soc.* **1988**, *110*, 8606. Thiem, H.-J.; Brandl, M.; Breslow, R. *J. Am. Chem. Soc.* **1988**, *110*, 8612.

NMR Template Analysis of Biphenomycin: The Prediction of Conformational Domains Defined by Clustered Distance Constraints

Judith C. Hempel*,† and Frank K. Brown

Contribution from the Department of Physical and Structural Chemistry, Smith Kline and French Laboratories, P.O. Box 1539, L950, King of Prussia, Pennsylvania 19406-0939.
Received February 1, 1989

Abstract: The stereochemistry of the cyclic tripeptide antibiotic biphenomycin has been assigned by using conformational information derived from NMR studies. Kannan and Williams[1] argue from a smaller and less quantitative set of NMR observations in the assignment of stereochemistry than do Brown and co-workers.[2] Energy-based modeling studies[2] suggest that both data sets are sufficient to assign the chirality of residue 3 and demonstrate that the smaller data set cannot be used to assign chirality at residue 1. Template analyses[3,4] of the two data sets are reported in this study. This analysis predicts conformational domains defined by sets of clustered NMR observations and identifies regions of the molecule where conformational variability associated with unconstrained rotatable bonds or "links" joining conformational domains can be expected in modeling studies. NMR template analysis is an ab initio analysis of the distance constraints and is not derived from the results of modeling studies. The results of the energy-based modeling analysis of the two data sets reported by Brown et al.[2] are consistent with the predictions of the template analysis.

Biphenomycin A (Figure 1) is a cyclic tripeptide antibiotic that shows potent antibacterial activity against gram-positive organisms.[5] The stereochemistries of residues 1 (atom 14) and 3 (atoms 8 and 7) have been assigned by using a combination of NMR and computational techniques.[1,2] The chirality of residue 2 (atom 11) was assigned independently.[5] All three residues are characterized by *S* chirality at the α carbon, while the β carbon of residue 3 is *R* (atom 7). Two groups of workers undertook the assignment of the chiral centers of residues 1 and 3 using con-

† Present address: Biosyn Technologies, Inc., 10065 Barnes Canyon Rd., Suite H, San Diego, CA 92121.

(1) Kannan, R.; Williams, D. H. *J. Org. Chem.* **1987**, *52*, 5435–5437.
(2) Brown, F. K.; Hempel, J. C.; Dixon, J. S.; Amato, S.; Mueller, L.; Jeffs, P. W. *J. Am. Chem. Soc.*, following paper in this issue.

0002-7863/89/1511-7323$01.50/0

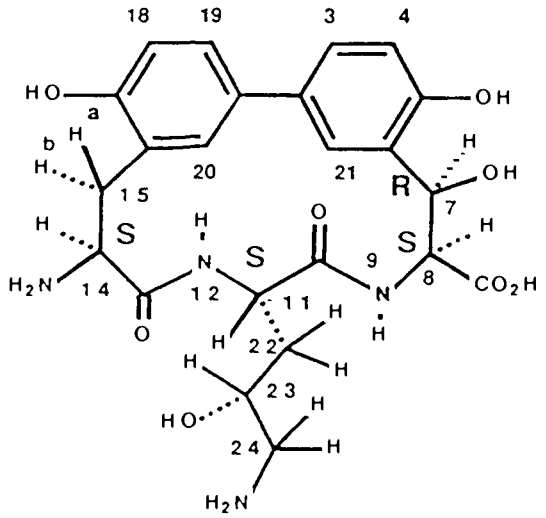
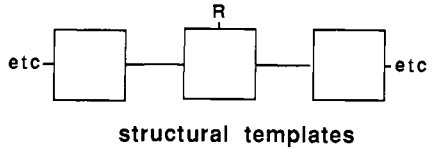
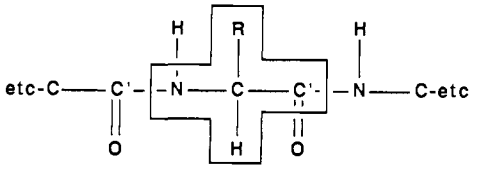


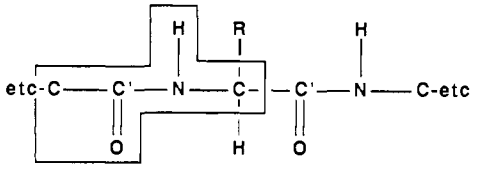
Figure 1. Structure of the cyclic tripeptide antibiotic biphenomycin A. Chiral centers are assigned.



structural templates



sp3 carbon and attached atoms



peptide bonds

Figure 2. Structural templates. The rotatable bonds of a molecule may be defined by solid lines connecting structural templates. Two types of structural templates associated with the peptide backbone are defined in this example.

formational information derived from NMR studies.[1,2]

Template analysis is a systematic procedure for mapping the distribution and correlation of the NMR-derived distance constraints across the molecule of interest.[3,4] When observations are correlated, the value chosen for one constraint places limits on the allowed values for all others. Conformational domains are defined by sets of correlated constraints. An NMR template analysis identifies the conformational domains defined by correlated constraints and predicts where conformational variability associated with unconstrained rotatable bonds and "linker" regions joining conformational domains can be expected in modeling studies. The template analysis is an ab initio analysis of the distance constraints and is not derived from the results of modeling studies. Template analyses of the two data sets used in the assignment of stereochemistry for biphenomycin are presented in this study. The analyses of the data sets are illustrated with molecular models built[2] to express the data set.

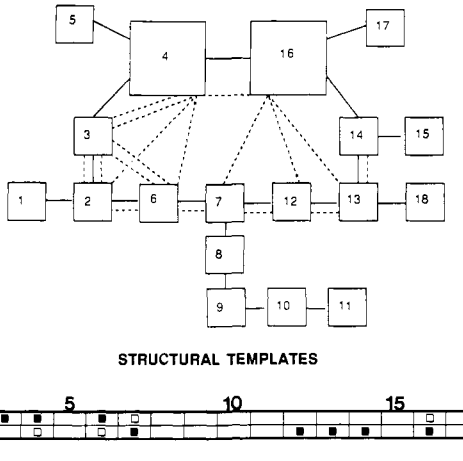


Figure 3. Data set 1. Top: the rotatable bonds of the biphenomycin molecule are defined by solid lines connecting structural templates. Dotted lines represent NOEs reported by Brown et al.[2] Bottom: the rows of this table define conformational domains. Columns index the structural templates of the molecule. Filled squares identify the structural templates that define conformational domains. Open squares identify pendant templates that are linked to the domains by isolated data.

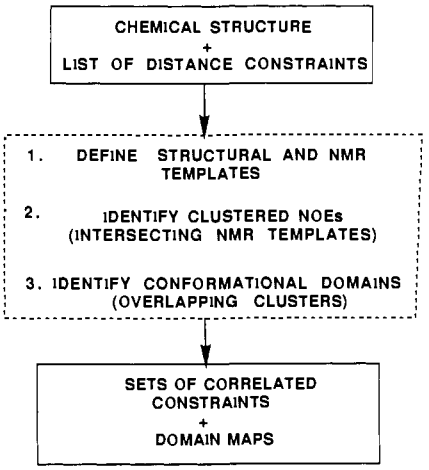


Figure 4. Domain mapping strategy. Analysis algorithms divide the list of distance constraints into sets of correlated constraints and map the conformational domains defined by the distance constraints. (1) Structural templates are sets of atoms for which all interatomic distances are specified by bonding arguments. NMR templates are sets of atoms for which all interatomic distances are specified by bonds and/or NOEs. (2) All NMR templates that share an NOE add to the definition of a set of correlated NOEs. Structural templates linked by these clustered NOEs map the region of the molecule that is conformationally constrained by the NOEs. (3) Conformational domains are defined by the union of all clusters that overlap by two or more structural templates.

Constraint Analysis

The predictions of the NMR template analysis derive from the fact that the conformational constraints that define the data sets under analysis are not uniformly distributed throughout the molecule. It is informative for the purposes of this discussion to represent the rotatable bonds of biphenomycin as solid lines connecting structural templates and to add the conformational constraints derived from NMR studies as dotted lines. Structural templates are defined by sets of atoms whose internal conformation

(3) Hempel, J. C.; Heald, S. L.; Jeffs, P. W. *Proceedings of the 10th Americal Peptide Symposium*; Marshall, G. R., Ed.; ESCOM Science Publishers: B. V., The Netherlands, 1988; pp 62–64.

(4) Hempel, J. C. An NMR Distance Constraint Analysis for the Aglycon of Aridicin A Bound to Ac-Lys(Ac)-D-Ala-D-Ala: A New Method for Predicting Well-Defined and Variable Regions of Molecular Conformation Defined by Distance Constraints. *J. Am. Chem. Soc.* **1989**, *111*, 491–495.

(5) (a) Uchida, I.; Ezaki, M.; Shigematsu, N.; Hashimoto, M., *J. Org. Chem.* **1985**, *50*, 1341. (b) Ezaki, M.; Iwami, M.; Yashita, M.; Kohsaka, M.; Aoiki, H.; Imanaka, H. *J. Antibiol.* **1985**, *38*, 1453. (c) Uchida, I.; Shegimatsu, N.; Ezaki, M.; Hashimoto, M. *J. Antiobiot.* **1985**, *38*, 1462.

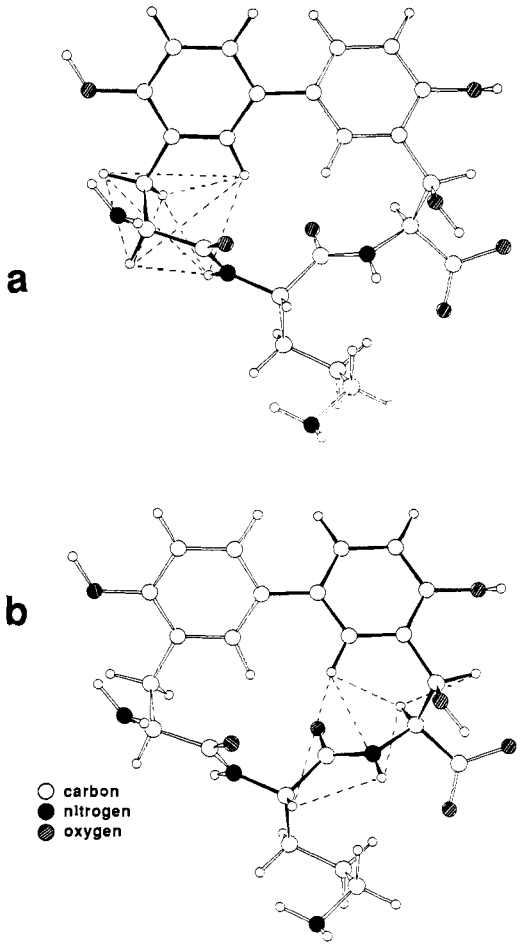


Figure 5. Conformational domain models. (a) Row 1 of Figure 3 (templates 2-3-4-6). Note that the NOEs (dotted lines) that define this conformational domain define an NMR template. (b) Row 2 of Figure 3 (templates 7-12-13-14-16). Note that NOEs that define the conformational domain define intersecting NMR templates. The domain includes a bridging structural template (template 14) that is bonded to two structural templates mapped by intersecting NMR templates. The NOE linking the α and β hydrogens of residue 3 (templates 13 and 14) is correlated across the domain for this reason. Bridging templates are identified by the analysis algorithms.[4,6]

is fixed by chemical bonding arguments. Structural templates associated with the peptide backbone are defined in Figure 2. Note that peptide bonds are treated as a single structural template. This assignment is equivalent to an assumption that the peptide bond is either cis or trans but is not converting between these two geometries. The peptide bonds of the biphenomycin molecule were assigned trans in the data sets under analysis.[1,2]

Data Set 1 (Brown et al.). Nuclear Overhauser effects (NOEs) determined by Brown et al.[2] are illustrated schematically in Figure 3 as dotted lines connecting structural templates. The NOEs in this data set were quantitated by monitoring buildup rates in 2D NOE studies. Coupling constants provide redundant information at relevant torsion angles. Distance constraints associated with unobserved NOEs were assigned but are not included in the analysis. The clustering of the NOEs in the data set is evident from this representation of the data. The identity of the structural templates of biphenomycin is established by a comparison of Figures 1 and 3. Conformational domains defined by correlated NOEs are mapped by using a table format in which the columns of the table index the structural templates of the molecule.

The domain map presented in Figure 3 is generated algorithmically in a three-step process illustrated schematically in Figure 4. First, NMR templates are defined. These are sets of atoms for which all interatomic distances are specified by bonding and/or NOEs. Second, clusters of correlated NOEs are identified. Each NMR template that shares one or more NOEs with another adds

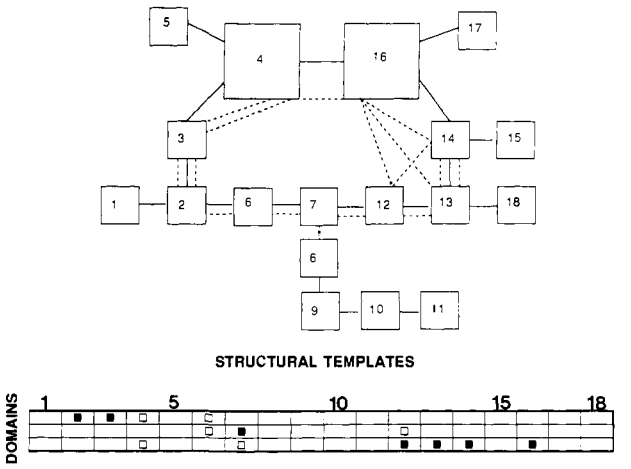


Figure 6. Data set 2. Top: rotatable bonds of the biphenomycin molecule are indicated by solid lines connecting structural templates. Dotted lines represent the NMR observations used by Kannan and Williams[1] in an assignment of stereochemistry for this molecule. Bottom: row 3 of the table defines a conformational domain (four bonded structural templates linked by clustered data). Open squares (row 3) identify pendant templates that are linked to the domain by isolated data. Rows 1 and 2 identify rotatable bonds constrained by isolated data.

to the definition of a cluster of correlated NOEs. The structural templates that are linked by correlated NOEs map a constrained cluster region of the molecule. Third, conformational domains are defined by the union of all clusters of structural templates that overlap by two or more structural templates. The analysis algorithms used in the definition of the domain map have been defined previously.[3,4] The algorithms require only that the chemical structure of the molecule be defined and the identity of the conformational constraints be specified.

The full set of distance constraints is mapped. In this example, not all distance constraints contribute to the definition of a conformational domain (i.e., three or more structural templates linked by correlated data). These extra constraints are called isolated constraints. Filled squares are used in the domain map defined in Figure 3 to map structural templates constrained by correlated NOEs (i.e., domains). Open squares identify structural templates that are pendant to a conformational domain. Pendant templates are templates that are linked to only one structural template of the domain by bonds and/or isolated data. Two NOEs in this data set are not included in the definition of a conformational domain, i.e., they are isolated and define pendant templates.

The domain map reveals at a glance regions of the molecule that are not constrained by NOEs. These include the side-chain of residue 2 (templates 8-9-10-11) and five structural templates pendant to the molecule by bonds alone (templates 1, 5, 15, 17, 18). These templates correspond to the amino and carboxy terminus of the molecule and the hydroxyl substituents of residues 1 and 3.

Row 1 identifies a domain defined by the structural templates 2-3-4-6. Structural templates 7 and 16 are pendant to this domain. NOEs that define the domain are illustrated by dotted lines in Figure 5 by using a molecular model built to express the NOE distance constraints.[2] It is apparent from Figure 5 that the NOEs are correlated, viz., that once a distance is fixed for one NOE, that choice influences the allowed choices for all others. The conformational restrictions associated with correlated NOEs derive in part from geometrical requirements dictated by hybridization geometries and bonding rules. Once distances associated with the NOEs that define this domain are fixed in a modeling study, all interatomic distances associated with templates 2-3-4-6 have been defined. A consideration of the correlated NOEs that define the conformation of the domain defined by the structural templates 7-12-13-14-16 (row 2) that is also presented in Figure 5 further illustrates the concept of a conformational domain.

The conformational domains defined by the data set are linked by rotatable bonds defined by structural templates 4-16 and 6-7

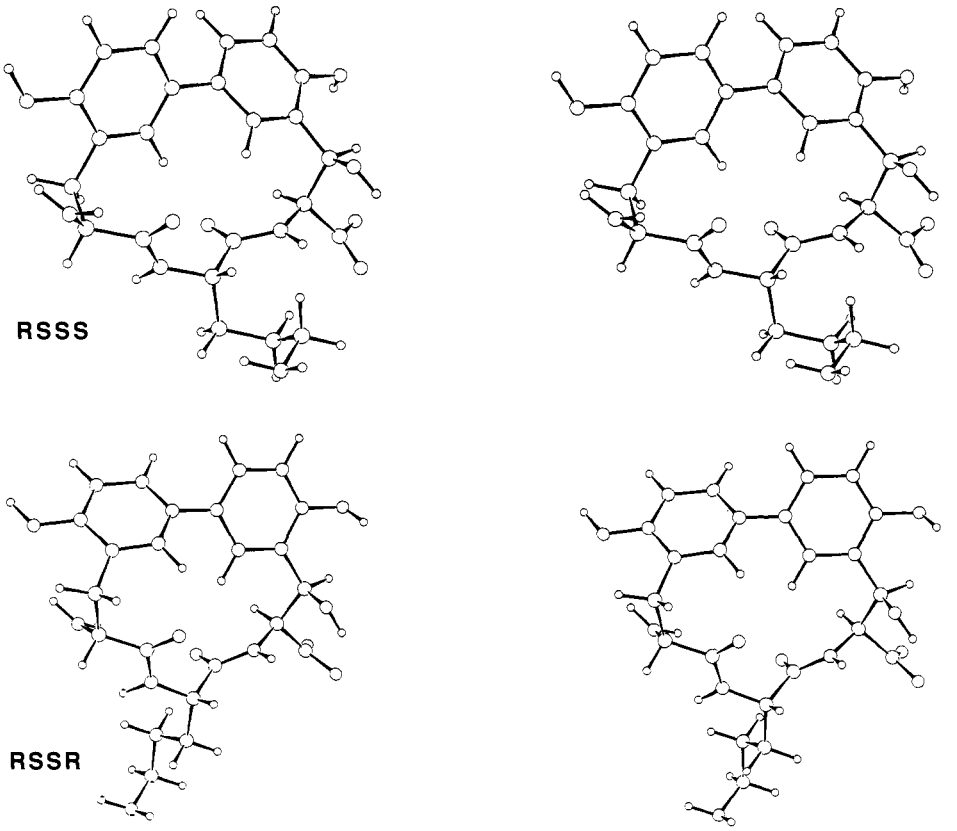


Figure 7. Low-energy conformations of biphenomycin with *RSSS* (top) and *RSSR* (bottom) stereochemistry that are consistent with the Kannan and Williams data set[1] presented in Figure 6. All torsion angles quantitated by Kannan and Williams are obeyed. NOEs observed by Kannan and Williams (but not quantitated) obey distance bounds defined by Brown et al.[2] This data set cannot be used to assign stereochemistry at residue 3. The models are presented as stereo pairs. Note that the conformation of the side chain of residue 2 is not defined by the NMR data set.

to form the ring of the cyclic tripeptide. Note that the linked domains are overlapped in the domain map. The overlap of structural templates pendant to one region (open square) and integral to the other (filled square) denotes that the rotatable bond involved (e.g., 4-16 or 6-7) is constrained by isolated data.

Data Set 2 (Kannan and Williams). The data used by Kannan and Williams[1] in their assignment of the chiral centers of biphenomycin is analyzed in Figure 6. As in Figure 5, solid lines represent rotatable bonds linking structural templates. However, in Figure 6 the dotted lines represent *J* coupling constraints, *qualitative* NOEs, and unobserved NOEs when the failure to observe an NOE was included in a decision leading to the assignment of stereochemistry.[1]

An inspection of the data set presented in Figure 6 reveals that each rotatable bond defined by the ring portion of biphenomycin is constrained. However, with the exception of the torsion angles defined by structural templates 12-13-14-16, the rotatable bonds are constrained by isolated data linking sequential structural templates. One conformational domain is defined in the domain map (identified by filled squares in row 3). Isolated constraints define structural templates pendant to the cluster (open squares in row 3), a sequence of three rotatable bonds constrained by isolated data (defined by templates 6-2-3-4 in row 1), and a sequence of two rotatable bonds (defined by three templates 6-7-12 in row 2). Rows 1, 2, and 3 overlap to define a series of six rotatable bonds constrained by isolated data (defined by templates 16-4-3-2-6-7-12). This region is linked to the conformational domain defined by correlated data (templates 12-13-14-16) by bonds 16-4 and 7-12.

Discussion

Kannan and Williams[1] argue from a smaller and less quantitative set of NMR observations in the assignment of stereochemistry for biphenomycin than do Brown and co-workers.[2] Although both groups reach the same conclusions (Figure 1), a comparison of the data sets presented in Figures 3 and 6 reveals that the NMR-derived conformational arguments in the smaller data set are clustered in such a way that conformational variability consistent with these arguments may "allow" either *R* or *S* chirality at residue 1.

Solution conformation models that satisfy the data set can be generated[2] with *RSSS* and *RSSR* stereochemistries (Figure 7). The *RSSS* nomenclature implies that the stereochemistry of the β carbon of residue 3 (atom 7) is *R* and that of the α carbons of residues 3, 2, and 1 (atoms 8, 11, and 14) is *S*. The models presented in Figure 7 are consistent with the torsion angles quantitated by Kannan and Williams. Interproton distances associated with NOEs observed by Kannan and Williams (but not quantitated) obey distance bounds defined by Brown et al.[2] An inspection of Figure 7 reveals how conformational compensation can occur without introducing significant strain into the molecule with this data set when the chirality at residue 1 is changed from *S* to *R*. By contrast, solution conformation models[2] with alternate stereochemical assignments at residue 3 are appreciably strained. The energy-based modeling studies[2] demonstrate that the stereochemistry at residue 1 cannot be assigned with this data set but suggest that the stereochemistry at residue 3 can.

It has been previously hypothesized[4,6,7] that when distance constraints are correlated they need not be precisely determined to define conformation. For example, Kannan and Williams[1] correctly assign the stereochemistry at residue 3 using qualitative (but *correlated*) conformational information from their data set and hand-held models. The hypothesis is also supported[4,6] by an analysis of a solution conformation model of the aglycon of the glycopeptide antibiotic Aridicin A (a heptapeptide) bound to Ac2-L-Lys-D-Ala-D-Ala. The model was defined by using NOEs

(6) Mueller, L.; Heald, S. L.; Hempel, J. C.; Jeffs, P. W. Determination of the Conformation of Molecular Complexes of the Aridicin Aglycon with Ac2-L-Lys-D-Ala-D-Ala and with Ac-L-Ala-(gamma)-D-Gln-L-Lys-(Ac)-D-Ala-D-Ala: An Application of Nuclear Magnetic Resonance Spectroscopy and Distance Geometry in the Modelling of Peptides. *J. Am. Chem. Soc.*, in press.

(7) Go, N.; Scheraga, H. A. *Macromolecules* **1970**, *3*, 178–187.

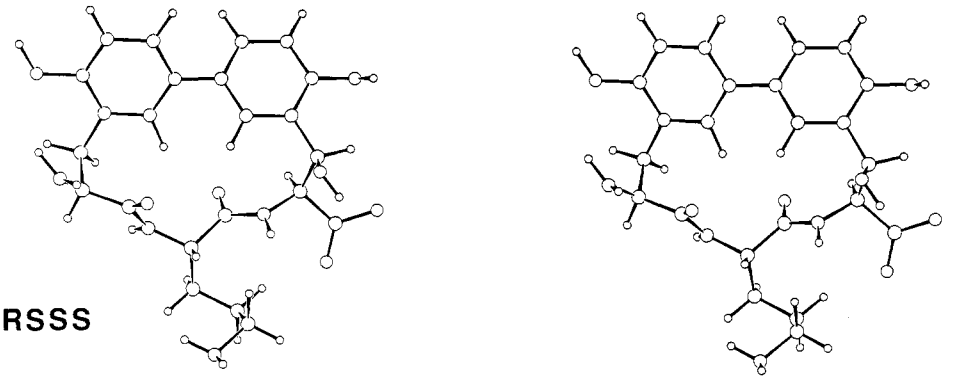


Figure 8. Solution conformation model of biphenomycin defined by Brown et al.[2] presented as a stereo pair. *RSSS* stereochemistry (Figure 1) was established in energy-based modeling studies[2] incorporating the NMR distance constraints reported by Brown et al.[2] and summarized in Figure 3. Note that the conformation of the side chain of residue 2 is not defined by the NMR data set.

that were not quantitated by using buildup rates. The distance bounds for the NOEs were therefore not as precisely defined as they could have been with further studies to monitor buildup rates. A template analysis of the data set (107 constraints) reveals how and where the distance constraints that define the model are correlated and correctly predicts the existence of well-defined conformational domains that are observed in modeling studies.[6]

Figure 3 reveals that the NOEs reported by Brown et al.[2] define two conformational domains centered around residue 1 and residues 2-3. An *RSSS* stereochemical assignment (Figure 1) is the only stereochemical assignment allowed by this data set. The solution conformation model proposed for this molecule[2] is illustrated in Figure 8. Solution conformation models with other stereochemical assignments[2] are significantly strained (i.e., are characterized by significantly higher strain energy than the *RSSS* model) and, in some cases, violate torsion angle constraints implied by *J* coupling constants. The energy-based modeling studies that support these conclusions are reported in detail by Brown et al.[2]

Conclusion

An analysis technique[4,6] that integrates structural information with NMR-derived distance constraints is used to map regions of biphenomycin where the conformation is defined by correlated distance constraints. These conformational domains are defined by structural templates linked by clustered data. In general, the more highly correlated the NOEs, the more constrained the conformation of the molecular domain. Highly correlated and precisely determined NOEs have been used to assign the stereochemistry of biphenomycin and define the solution conformation of this molecule on the basis of energy-based modeling.[2] The assignment of absolute stereochemistry is made possible by an independent assignment of the chirality of residue 2.[5]

A variety of molecular modeling techniques including constrained molecular dynamics, distance geometry, and systematic search are currently under investigation in a number of laboratories for probing the conformation space of a molecule subject to distance constraints.[8-13] It has been previously demonstrated that well-defined and variable regions can be identified in a statistical analysis of multiple conformers of a protein defined by interproton distances[9] and that regions of polypeptide chain that are the most variable between different structures typically have the lowest density of constraints and vice versa.[13] NMR template analysis is an ab initio analysis of the distance constraints that *identifies* the regions of a molecule constrained by correlated distance constraints.

The template analysis algorithms are simple[4] and may be readily applied to small proteins.[14] The analysis makes it possible to assign confidence levels for features of a protein model built to express interproton distance constraints. The origin of structural "errors" in these models can be traced by understanding the interdependence of the distance constraints, i.e., how the distance constraints that define that region of the model are correlated across the molecule. Template analysis is computationally cheap and can be done before modeling studies to facilitate error checking. This procedure not only highlights simple errors in data input but can help in the identification of distance bounds that contain no conformational information (e.g., when bounds are set for an interproton distance internal to a structural template). The analysis does not require that the distance constraints under investigation be derived from NMR studies but is a general method for mapping the conformational implications of clustered distance constraints.

(8) Fesik, S. W.; O'Donnell, T. J.; Gampe, R. T., Jr.; Olejniczak, E. T. *J. Am. Chem. Soc.* **1986**, *108*, 3165-3170.
(9) Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1987**, *219*, 11-16.
(10) Havel, T. F.; Wuthrich, K. *J. Mol. Biol.* **1985**, *182*, 281-294.
(11) Kaptein, R.; Zuiderweg, E. R. P.; Scheek, R. M.; Boelens, R.; Van Gunsteren, W. F. *J. Mol. Biol.* **1985**, *182*, 179-182.
(12) Clore, G. M.; Brunger, A. T.; Karplus, M.; Gronenborn, A. M. *J. Mol. Biol.* **1986**, *191*, 523-551.
(13) Wagner, G.; Braun, W.; Havel, T. F.; Schaumann, T.; Go, N.; Wuthrich, K. *J. Mol. Biol.* **1987**, *196*, 611-639.
(14) Hempel, J. C., unpublished results.