the 1:1 ratio of the two heme orientations formed during the initial contact (Figure 5E). However, it appears to act to diminish some of the strong preference in the heme orientation for the metal complex for the initial propionate contact for 6-methyl-6-despropionate-hemin (compare trace B and A of Figure 3) and 7-methyl-7-despropionate-hemin (compare traces B and A of Figure 4).

Structure of desFe Mb. The typical upfield ring-current shifted Val E11 γ-CH_3 peak in desFe Mb[19] confirms a heme pocket structure similar to that of MbCO.[24] However, both the upfield Val CH_3 and low-field meso-H ring current shifts are significantly larger than in MbCO[24] (compare in Table I). In fact, the low-field meso-H shifts are larger in desFe Mb than either MbCO[24] model complexes or free protoporphyrins,[38,39] which exhibit mean meso-H shifts ~9.8 ppm. The mean meso-H shift (10.3 ppm) in desFe Mb, however, is in the direction of that recorded for protoporphyrin IX in trifluoroacetic acid,[40] in which it exists as a dication, with mean meso-H shift ~11.0 ppm. Such dication formation has been shown to lead to substantial increase in the ring current due to resonance stabilization.[39-41] The ~40% of the low-field bias toward the dication suggests an intermediate state of protonation. Thus both the low-field bias of meso-H and upfield bias of Val E11 CH_3 in desFe Mb relative to MbCO strongly suggests protonation of the prosthetic group. While such an effect may not be expected in a presumed relatively nonpolar heme pocket, it could arise from strong hydrogen bonding between the protonated imidazolium side chains of the proximal His F8 and a porphyrin ring imine nitrogen. The structure of the complex is under further study.

(38) Caughey, W. S.; Barlow, C. H.; O'Keeffe, D. H.; O'Toole, M. C. *Ann. NY Acad. Sci.* **1973**, *200*, 296-308.
(39) Janson, T. R.; Katz, S. S. *The Porphyrins* **1978**, *4B*, 1-59.
(40) Abraham, R. J.; Jackson, A. H.; Kenner, G. W. *J. Chem. Soc.* **1961**, *3*, 3468-3474.
(41) Abraham, R. *J. Mol. Phys.* **1961**, *4*, 145-152.

Conclusions

The formation of holo-Mb with a 1:1 mixture of hemin rotationally disordered about the α,γ-meso axis in the initial complex formed during the reconstitution of Mb is due to the formation of a relatively stable salt bridge between a single propionate group and a unique protein residue, possibly Arg CD3. Removal of one of the propionate groups leads to a strong preference for one of the orientations that places the sole propionate into the position occupied by the 6-propionate group in the Mb X-ray structure. The pyridine solvent used to solubilize the hemin for reconstitution was found to interact specifically with the protein complex on the proximal side of the heme.

Acknowledgment. The authors are indebted for experimental assistance to the staff of the UCD NMR Facility. This research was supported by grants from the National Institutes of Health (HL-16087 and HL-22252).

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

Judith C. Hempel

Contribution from the Department of Physical and Structural Chemistry, Smith Kline and French Laboratories, Swedeland, Pennsylvania 19479. *Received May 2, 1988*

Abstract: Over 100 NMR-derived distance constraints are reported in a companion paper[1] for the aglycon of Aridicin A bound to the tripeptide Ac-Lys(Ac)-D-Ala-D-Ala. Even though the distance constraints are not as precisely determined as chemical bonds, correlated NOEs severely constrain the conformation of clusters of structural templates of the molecular complex. Overlapping clusters extend the regions of restricted conformation. A methodology is introduced to map the clusters of structural templates defined by the data set. The map reveals the regions of the molecular complex where the conformation is defined and regions where it is not. The more highly correlated the distance constraints, the more restricted the conformation of the molecular domain. The relevance of this ab initio analysis of the distance constraints is demonstrated with an analysis of seven molecular models for the solution conformation of the molecular complex generated by distance geometry. Conformational domains predicted by the ab initio analysis are observed in the molecular modeling study.

Aridicin A is a member of the vancomycin/ristocetin class of glycopeptide antibiotics. The aglycon of Aridicin A is a heptapeptide in which the aromatic side chains of residues 1 and 3 are linked, as are the side chains of residues 2 and 4, 4 and 6, and 5 and 7 (see Figure 1). The seven residues are named GCFBEAD proceeding from the N to the C terminus of the peptide in agreement with a previously defined convention.[2] The action of this class of antibiotics is postulated to involve binding to D-alanyl-D-alanine residues at the C terminus of bacterial cell wall components.[3] NMR studies are reported in a companion paper[1] for complexes of the aglycon of Aridicin A bound to the bacterial cell wall fragments Ac-Lys(Ac)-D-Ala-D-Ala and Ac-Ala-γ-D-Gln-Lys(Ac)-D-Ala-D-Ala. NMR-derived distance constraints are the same for the region common to both complexes. A single solution conformation model that satisfies all NMR-derived distance constraints for the molecular complex with tripeptide has been proposed.[1]

(1) Mueller, L.; Heald, S. L.; Hempel, J. C.; Jeffs, P. W. *J. Am. Chem. Soc.*, following paper in this issue.
(2) Jeffs, P. W.; Mueller, L.; DeBrosse, C.; Heald, S. L.; Fisher, R. *J. Am. Chem. Soc.* **1986**, *108*, 3063-3075.
(3) (a) Reynolds, P. E. *Biochim. Biophys. Acta* **1961**, *52*, 403-405. (b) Jordan, D. C. *Biochim. Biophys. Res. Commun.* **1961**, *6*, 16-170. (c) Perkins, H. R. *Biochem. J.* **1969**, *111*, 195-205.

0002-7863/89/1511-0491$01.50/0

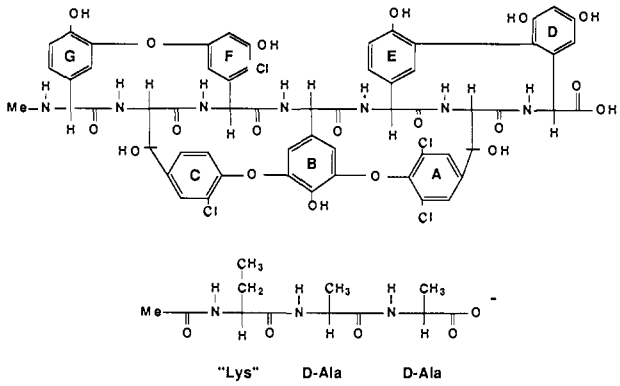

Figure 1. Chemical structure of aridicin A aglycon (top) and tripeptide Ac-Lys(Ac)-D-Ala-D-Ala (bottom). The residues of the aglycon are named GCFBEAD proceeding from the N to the C terminus of the heptapeptide. The Lys side chain of the tripeptide is truncated for the purposes of this analysis.

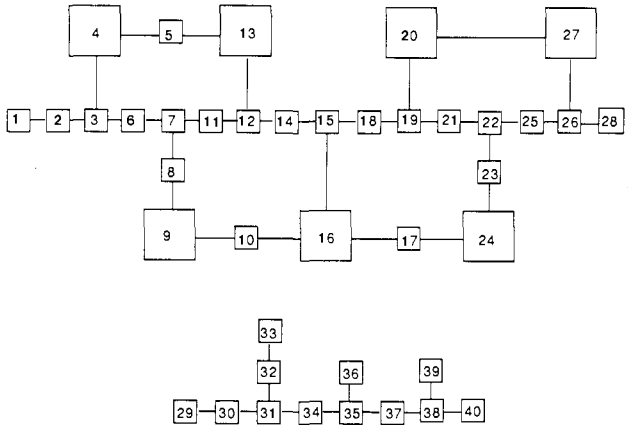

Figure 2. Structural templates. The rotatable bonds of the aglycon and tripeptide (Figure 1) are defined by the solid lines that connect boxes representing structural templates. The numbering scheme for the structural templates established in this figure is maintained in Figures 3 and 5. Structural templates associated with hydroxyl substituents are not included in the analysis.

The rotatable bonds of the core region of the molecular complex are defined in Figure 2 as solid lines connecting structural templates. As is implied by a comparison of Figures 1 and 2, structural templates are sets of atoms whose conformation is fixed by chemical-bonding arguments. A total of 107 NOEs and hydrogen-bonding interactions limit the degrees of freedom of this molecular complex and must be simultaneously satisfied in a solution conformation model. The identification of conformational domains, sets of structural templates whose conformation is constrained by correlated constraints, is discussed in this paper. In general, the larger the number of correlated distance constraints, the more constrained the conformation of the conformational domain. When distance constraints are correlated, the value chosen for one distance places limits on the allowed values for all others. When distance constraints are not correlated, all combinations of distances within the limits specified are, in principle, allowed.

A mapping algorithm is introduced that maps the clusters of structural templates linked by NOEs. Conformational domains are defined by interlocking clusters. The definition and analysis of the cluster map is discussed in detail in this paper. The map reveals how the distance constraints of the data set are correlated and identifies conformational domains defined by the data set. An analysis of seven molecular models for the solution conformation of the molecular complex previously generated[1] by distance geometry[4,5] is also reported. This analysis reveals that the variability in the molecular models is associated with "hinge" regions where conformational domains predicted by the ab initio analysis of the distance constraints are joined.

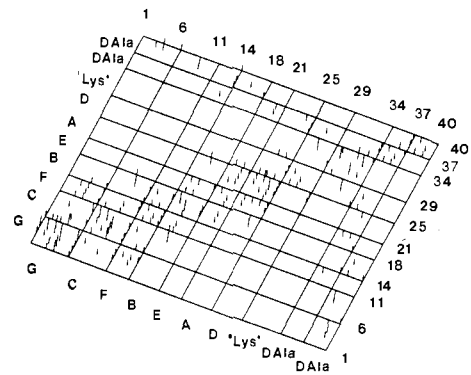

Figure 3. Distribution of NOE observations. Peak heights are proportional to the number of NMR observations in the data set linking the structural templates indexed by peak position. Grid blocks map the residues of the aglycon and tripeptide (Figure 2). Blocks are identified by residue name. Peptide bond templates are indexed by number.

Table I. Analysis of the Intersections of Three NMR Templates Defined in Figure 4[a]

NMR template	atoms	NOEs	no. of shared distances (NOEs)		
			b	*c*	*d*
b	4	1	6 (1)	3 (1)	3 (1)
c	5	5	3 (1)	10 (5)	6 (3)
d	5	6	3 (1)	6 (3)	10 (6)

[a] Atom names are established in Figure 4 and ref 1. [b] GMe, GN, GCA, GH1′. [c] GMe, GCA, GH1′, GH5, GH6. [d] GMe, GCA, GH1′, GH6, CNH.

Distance Constraint Analysis

Automated procedures have been developed that identify the structural templates of the molecule and map the clusters of structural templates linked by NMR distance constraints. The analysis requires that the chemical structure of the molecule be specified and the distance constraints be identified. The analysis provides a list of the distance constraints that define each cluster along with lists of the structural and NMR templates associated with each cluster. NMR templates are discussed below. The analysis of the NMR-derived distance constraints for the molecular complex of aglycon with tripeptide is outlined in this section. The algorithms are defined in the Appendix.

NMR Distance Constraints. The distribution of the NMR distance constraints over the molecular complex is summarized in Figure 3. The distance constraints are tabulated in ref 1. The height of each peak is proportional to the number of NOEs and/or hydrogen-bonding assignments that link the structural templates indexed by that peak position. The grid of the figure is scored to index the peptide bond templates of the aglycon and tripeptide (Figures 1 and 2). The grid blocks map the structure of the molecular complex by residue. Blocks are identified by residue name, by using the GCFBEAD convention for the residues of the aglycon and a three-letter code for the residues of the bound tripeptide. No NMR observations link the ether linkages of the aglycon (templates 5, 10, and 17), the "dummy" side chain of Lys (template 33), or the C terminus of the aglycon (template 28) to others. With these exceptions, NMR distance constraints link all structural templates defined in Figure 2 to at least one other structural template. Hydrogen-bonding assignments provide conformational information for the carboxy terminus of the bound peptide (template 40).

NMR Templates. Clusters of structural templates linked by NMR distance constraints are identified by NMR templates. Just as 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 bonding and/or NMR-derived distance constraints.[6] All NOEs and hydrogen-bonding constraints that contribute to the definition of an NMR template are fully correlated. The distances linking the atoms of an NMR template obey

(4) Havel, T. F.; Kuntz, I. D.; Crippen, G. M. *Bull. Math. Biol.* **1983**, *45*, 665–720.
(5) Crippen, G. M.; Havel, T. F. *Acta Crystallogr.* **1978**, *34*, 282–284.
(6) Hempel, J. C.; Mueller, L.; Heald, S. L.; Jeffs, P. W. *Proceedings of the 10th American Peptide Symposium*; Marshall, G. R., Ed.; ESCOM Science Publishers B.V.: Leiden, The Netherlands, 1988; pp 62–64.

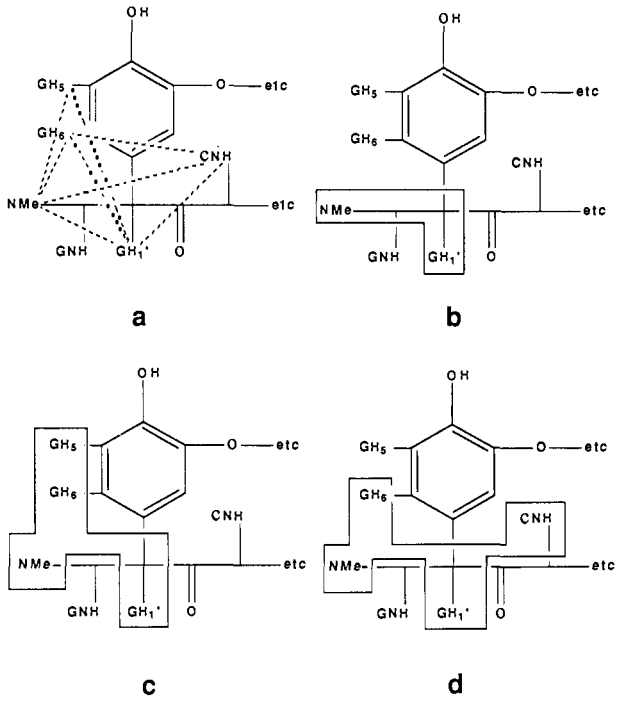

Figure 4. Three NMR templates. NMR distance constraints associated with the G residue of the aglycon (Figure 1) are summarized by dotted lines in (a). Three NMR templates defined by this data are defined by boxes in (b)–(d). This data defines the conformation of structural templates 1-2-3-4-6 (Figure 2) and contributes to the definition of clusters 1–4 in Figure 5. Atom names incorporate the one-letter residue code established in Figure 1 for aridicin A as the first letter of the atom name. NOE distance constraints are tabulated using this convention in ref 1.

all triangle (and higher) inequalities that can be defined for those atoms.

Three NMR templates linking templates of the G residue of the aglycon are defined in Figure 4. The NMR observations that define these NMR templates are indicated schematically in Figure 4a with dotted lines, and the atoms that define the three NMR templates are specified by boxes in the remaining panels of Figure 4. All atoms contributing to the NMR templates are included in the boxes.

Distance constraints imposed by NMR observations need not be precise to define conformation when the observations are correlated, as they are in this example. The correlation of the NMR-derived distance constraints defined in Figure 4a is revealed by an analysis of the NMR templates defined in Figure 4. The intersections of these templates are summarized in Table I. Each pair of NMR templates shares two or more atoms for which the interatomic distances are defined by an NMR distance constraint. Because the distances associated with the shared atoms are the same, the distances associated with the atoms that are not shared are correlated.

Cluster Map. Clusters of structural templates linked by NMR distance constraints are defined by sets of intersecting NMR templates where each NMR template shares two or more atoms with all others in that set. Clusters of structural templates defined in this way interlock like the pieces of a three-dimensional puzzle in molecular models built to express the NMR-derived distance constraints.

The identity of the clusters defined by the full set of NMR-derived distance constraints is presented as a cluster map in Figure 5. Columns index the structural templates defined in Figure 2. Each row of the table is defined by a different set of intersecting NMR templates. Redundancy in the cluster definitions implies redundancy in the data set. Filled squares identify the structural templates linked by correlated distance constraints. Unfilled squares distinguish structural templates linked to the cluster by NMR constraint(s) that are not correlated across the cluster. The conformation of these pendant templates relative to the cluster is defined only to within a mirror-plane reflection. Figure 5 reveals at a glance where the conformation of the molecule is constrained by correlated NOEs and where it is not.

The analysis summarized in Table I reveals that the distance constraints defined in Figure 4 contribute to the definition of a cluster since each of the NMR templates defined by these distance constraints shares two or more atoms. The three NMR templates (Figure 4 and Table I) are defined by a set of correlated distance constraints and contribute to the cluster definitions presented in rows 1–4 of Figure 5 (i.e. define the filled squares in columns 1-2-3-4-6 of Figure 5). A comparison of Figures 2 and 4 is sufficient to identify structural templates 1-2-3-4-6 as structural templates whose relative conformation is restricted by the distance constraints defined in Figure 4.

Conformational Domains. Conformational domains are defined by the union of all clusters that overlap by two or more structural templates. Consideration of the redundancy of the cluster map (i.e. data set) may suggest division of the conformational domains defined in this fashion into smaller domains. Conformational domains are regions of the molecule where the conformation is defined by correlated distance constraints.

The distance constraints that define each cluster are correlated across that cluster and across all clusters that overlap by two or more structural templates. All structural templates linked by an NMR template are linked by NOEs (or hydrogen-bonding constraints) or bridge two structural templates linked by NMR-defined distance constraints. Therefore, all structural templates of the cluster are linked by bonding and/or NMR-derived distance constraints to two (or more) other structural templates of that cluster. The pendant structural templates distinguished by unfilled squares in Figure 5 are linked by bonding and/or distance constraints to one structural template of the cluster. Pendant templates defined in Figure 5 that are not overlapped in the definition of a conformational domain are pendant to the conformational domain.

Results

Distance Constraint Analysis. An inspection of the cluster map for the molecular complex presented in Figure 5 reveals that overlapping clusters constrain the conformation of templates 1–25 and templates 25–27 of the aglycon. However, while the conformation of the GCFB region of the aglycon (templates 1–16) is defined by multiple interlocking clusters, these clusters overlap by only two templates (templates 15 and 16) with the clusters of the BEA region of the molecule (templates 15–25). It is apparent

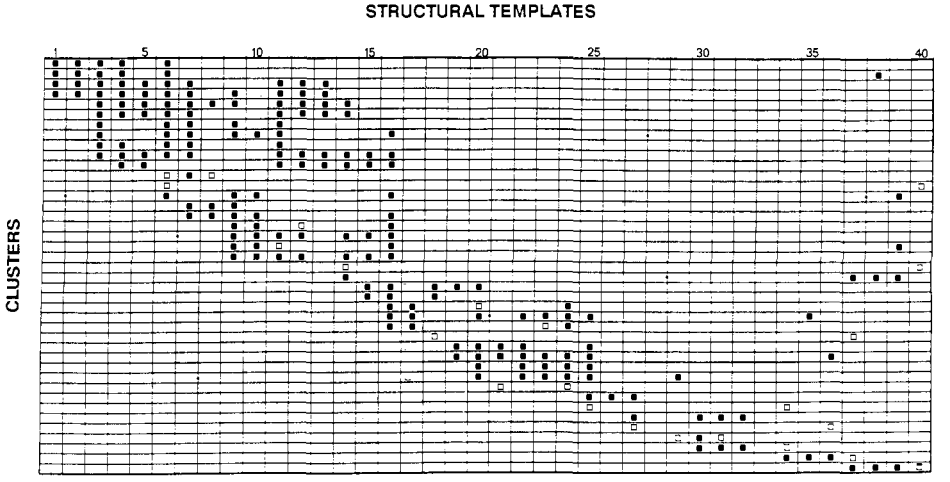

Figure 5. Cluster map. Clusters of structural templates (Figure 2) held constrained by NMR-derived distance constraints (Figure 3) are identified by row. Structural templates that are pendant to a cluster are distinguished with unfilled squares.

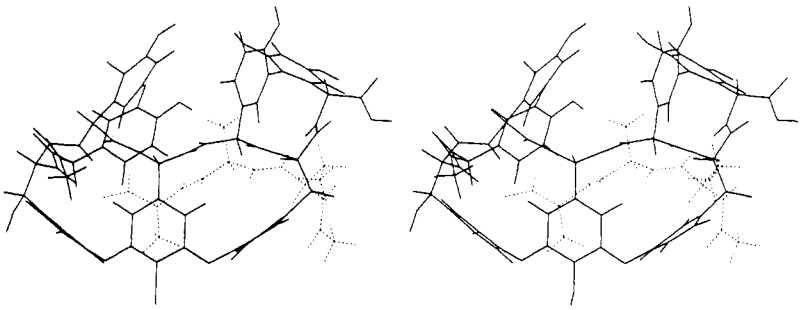

Figure 6. Solution conformation model for the molecular complex. The model is illustrated as a stereo pair. The bound tripeptide is distinguished with dotted lines.

that small changes in the side-chain orientation of residue B (defined by templates 15 and 16) can produce significant changes in the relative orientation of the GCF and EA regions of the aglycon with this data set. For this reason, two conformational domains are assigned in the GCFBEA region of the molecule.

The cluster map predicts that the GCFB region of the aglycon will be the most conserved domain of the aglycon in modeling studies because the conformational information in this region of the molecule is the most redundant of the data set. As outlined above, it is predicted that small changes in the side-chain orientation of residue B (defined by templates 15 and 16) can produce significant changes in the relative orientation of the GCFB and BEA regions of the aglycon. Limited conformational variability is expected in the BEAD region of the molecule where the BEA domain (templates 15–25) overlaps the D domain (templates 25–27).

There are always four conformational possibilities (in principle) when domains are joined by overlap of a single structural template because distance constraints do not distinguish between cluster conformations related by a mirror plane. For molecules characterized by rings and chiral centers fewer conformational possibilities are usually allowed. When the BEA and D domains are joined, the ring formed by the side chains of residues E and D must be closed. This ring is made up of only seven structural templates. Not only is the ring small but the relative chirality of five chiral centers must be maintained when the conformational domains are joined. Five templates of the seven-template ring are constrained by the correlated NOEs of the BEA domain in this data set. Three are constrained by the correlated NOEs of the D domain.

Variability is expected in the conformation of the bound peptide. Three domains of the bound tripeptide are constrained by correlated intrapeptide distance constraints. These are defined (Figure 3) by templates 30-31-32 of Lys-3, templates 34-35-36 of D-Ala-2, and templates 37-38-39 of D-Ala-1. These domains do not overlap. Additional conformational restrictions are imposed through binding to the aglycon. However, it is instructive to consider the intrapeptide conformational restrictions of the bound tripeptide further. The peptide bond templates of the tripeptide (templates 34 and 37) are each integral to one conformational domain and designated pendant to another in Figure 5. Since distance constraints do not distinguish between conformations related by a mirror plane, there are always two conformations for a structural template linked to a cluster by distance constraint(s) that are not correlated across the cluster. In the absence of constraints imposed by the chiral centers and by binding to the aglycon, this leads to eight conformational possibilities at each of the regions where the conformational domains of the bound tripeptide are joined. Most of these possibilities are not allowed in the context of the molecular complex. However, conformational variation is to be expected in these regions of the tripeptide in modeling studies.

Conformational Analysis. Seven models for the solution conformation of the molecular complex were generated by distance geometry[4,5] in a companion study.[1] The chiral centers were defined. The models agree with the NMR distance constraints of the data set to within an average error of 0.024 Å (107 constraints). A pair-wise comparison of the conformers demonstrates that all conformers are the same to within a least-squares root-mean-square (rms) deviation of 0.68 Å (over 148 atoms, excluding methyl and hydroxy hydrogens and the carboxy terminus of the aglycon). The conformer most different by the rms test also has the largest NMR distance constraint error. Six of the conformers are very similar (pair-wise rms <0.4 Å), and three are virtually identical (pair-wise rms <0.2 Å).

Three families of conformations are observed for the bound tripeptide. These are defined by a pair-wise rms fit of <0.1 Å within each family (over the atoms of the tripeptide, excluding methyl hydrogens). The three families are the same to within a 25° deviation of a φ and/or ψ angle of the tripeptide. The conformational families vary in regions where the conformational domains of the tripeptide (identified above) are joined.

The cluster map defined by the NMR distance constraints predicts that the conformational variation observed in solution conformation models will be due in large part to differences in the relative orientation of the GCFB and BEAD regions of the aglycon. This is observed. The six most similar conformers generated in the distance geometry study are identical to within an rms deviation of 0.18 Å over the GCFB residues of the aglycon (69 atoms, excluding methyl and hydroxy hydrogens) and, with the exception of a conformer with a chirality error at the β-carbon of the A residue (template 23), to within an rms deviation of 0.17 Å over the BEAD residues (64 atoms, excluding hydroxy hydrogens and the carboxy terminus). The six conformers are identical in the BED region to within an rms deviation of 0.12 Å (28 atoms). The variation observed in the conformation of the aglycon in these models is therefore due in large measure to variations in the relative orientations of the two sections of the molecular complex.

Conformational variation was investigated further by rms comparisons over the clusters defined by the data set (i.e. the sets of structural templates identified by row in Figure 3). All seven conformers are identical from the N terminus (template 1) through the region of the side chain of the C residue of the aglycon (template 8). The conformer identified as "different" in the comparisons described above differs by a mirror image tilt of the C ring (template 9) relative to that of the other conformers. The different conformation deviates from the others in the regions of the FBEA residues of the aglycon. A comparison of the clusters that define the conformation of the D residue of the aglycon and the "Lys"-3 residue of the bound tripeptide reveals that the conformation of these structural templates is once again the same in all seven conformers. There is less conformational diversity in clusters originating in the GCFB region of the aglycon than in the clusters originating in the BEAD region (i.e. the average rms deviation over the atoms of these clusters is less).

Solution Conformation Model. A stereo figure of the solution conformation model proposed for the binding region of the molecular complex of the aglycon of aridicin A bound to tri- or pentapeptide is given in Figure 6. This model is an energy-minimized version[1] of the three identical conformers identified above. The conformation of the aglycon of aridicin A when complexed is more "folded" than the conformation previously proposed for the uncomplexed aglycon of aridicin A in solution.[2]

It is proposed, on the basis of the analyses reported here and the similarity of the molecular models generated in the distance geometry study, that the conformation of the aglycon is the same when bound to the pentapeptide as when bound to the tripeptide. These results support the proposal that the solution conformation of the aglycon of Aridicin A bound to the tripeptide Ac-Lys-(Ac)-D-Ala-D-Ala models the binding conformation of the antibiotic. The conformation proposed for the aglycon is similar to that previously proposed by Fesik et al.[7] for ristocetin bound to Ac-Lys(Ac)-D-Ala-D-Ala. The models are discussed in greater detail in the companion paper.[1]

Discussion

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.[8] The analysis requires that a set of models representative of the conformation space of the protein be available for analysis. By contrast, the ab initio distance constraint analysis introduced in this paper *predicts* regions where the conformation of a molecule is defined by a set of distance constraints and regions where it is not. The analysis of the NMR distance constraints reported for the molecular complex of the aglycon of aridicin A requires less than 1 CPU min of computer time (VAX 11/780). A total of 75 cpu min (VAX 11/780) was required in the companion study[1] to generate and refine a molecular model for the molecular complex by the distance geometry algorithm EMBD.[4,5]

Conclusions

An analysis technique has been introduced that integrates structural information with NMR-derived distance constraints to predict molecular domains where the conformation is defined by the NMR distance constraints. Conformational domains are defined by interlocking clusters of structural templates that are linked by NOEs. The more extensively correlated the NOEs, the more constrained the conformation of the molecular domain.

This analysis technique does not require that the distance constraints under investigation be derived from NMR studies but is a general technique for mapping the conformational implications of distance constraints. Applications can be envisioned in which cluster mapping is used to prioritize experimental studies by establishing which studies yield new conformational information for the molecule of interest and which yield redundant or confirmatory information.

Applications can be envisioned where cluster mapping is used to define modeling strategy. In the simplest possible application, a cluster analysis could be used to identify domains defined by sets of highly correlated distance constraints that are then modeled and treated as structural aggregates in modeling studies of the molecule as a whole. 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.[7-12]

When conformational error is encountered in a modeling study, an analysis of the error associated with each conformational cluster is particularly informative. Error localized in this fashion suggests that there are inappropriate distance constraints in the data set associated with a cluster (or set of interlocking clusters). When the distance constraints are derived from NMR measurements, this can suggest regions of conformational flexibility in the molecule. In practice, the cluster map often highlights simple errors in data input prepared for modeling algorithms and facilitates the checking of alternate NMR assignments in a data set.

Acknowledgment. The helpful comments of F. K. Brown, P. W. Jeffs, K. D. Kopple, and I. D. Kuntz and the assistance of D. L. Brown and S. W. Lindsay in the preparation of the figures are gratefully acknowledged.

Appendix

Cluster Mapping Algorithms. 1. The *template* algorithm searches an $n \times n$ matrix in which nonzero entries indicate that the interatomic distance for the atoms indexed by row and column position have been specified. The algorithm checks systematically by rows to identify the distances *ij*, *ik*, ... that have been specified. When distances *jk*, ... are also specified, then atoms *i*, *j*, *k*, ... are assigned to a template. Structural templates are defined by distances derived from structural arguments alone. NMR templates incorporate an NMR-derived distance constraint.

2. The *cluster* algorithm defines a matrix that contains the intersections of the templates defined by the template analysis. The algorithm checks the intersections of the NMR templates systematically by rows to identify pairs of NMR templates *ij*, *ik*, ... that share two or more atoms. When pairs of NMR templates *jk*, ... also share two or more atoms, then the NMR templates *i*, *j*, *k*, ... can be used to identify a cluster of structural templates whose relative conformation is constrained by NMR-derived distance constraints. The cluster of structural templates is identified by the intersections of NMR and structural templates. Structural templates bridging two structural templates whose relative conformation is defined by the set of intersecting NMR templates are included in the cluster definition. Bridging templates are identified from the bonding pattern of the structural templates revealed by the intersections of structural templates with structural templates. Bonded structural templates share two atoms.

3. The *correlation of the distance constraints* associated with each cluster is traced by the template intersection matrix. The distance constraints that define the cluster are correlated across all structural templates connected by bonding and/or NMR-derived distance constraints to two other structural templates of the cluster. The template intersection matrix is used to identify structural templates that are pendant to a cluster (i.e. connected by bonding and/or NMR distance constraints to one other structural template of the cluster). The intersections of each structural template of the cluster with the other structural templates of the cluster reveal the structural templates that are bonded. The intersections of the NMR templates that define the cluster with the structural templates of the cluster reveal the additional connections introduced by the NMR-derived distance constraints. The correlation of the distance constraints across clusters is traced by identifying all clusters that overlap by two (or more) structural templates.

4. *Conformational domains* are defined by the union of all clusters that overlap by two or more structural templates. Structural templates pendant to a conformatinal domain are identified whenever templates pendant to a cluster are not overlapped in the definition of the conformational domain.

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