Abstract Submission 2016

Lognumber: 323 **Submitted on:** 1/21/2016 **Title:** Redefining Structure: New Approaches for Deep Topological Sampling of Protein Structures, and Continuous Probability Density Function for Biomolecular Interdomain Orientations

- Correspondent
- Abstract Entry
- Long Abstract Upload
- Authors and Affiliations
- Session Selection
- Additional Information (Req)
- Review and Submit

Review and Submit

Review your abstract for accuracy. Click **"Submit" or "Save"** when you are done.
Once submitted an email confirmation will be sent to all authors.

Title

Redefining Structure: New Approaches for Deep Topological Sampling of Protein Structures, and Continuous Probability Density Function for Biomolecular Interdomain Orientations

Authors
Yang Qi[1]; Jeff Martin[1]; Francois Thelot[1]; Adam Barb[2]; Tony Yan[1]; Pei Zhou[1]; Terry Oas[1]; Bruce Donald[1]
[1]Duke University, Durham, North Carolina; [2]Iowa State University, Ames, IA

Abstract
I will describe two novel approaches that redefine the concept of structure from NMR data. The first approach samples fold space using a topological group action on conformation space. It is useful for automated assignment and structure determination of symmetric homoöligomers with ambiguous distance restraints. In the second, we developed a new method for determining the structure of a flexible protein by RDCs that avoids the use of over-parameterized conformational ensembles. Our approach describes the interdomain orientation as a continuous 3D probability distribution, exploiting new geometry, a fitting algorithm, and a new graphical representation. The algorithms are demonstrated, respectively, to search fold space for the trimeric membrane protein Diacylglycerol Kinase (DAGK), and elucidate inter-domain dynamics of staphylococcal protein A (SpA).
View Long Abstract... (/portals/0/abstracts2016/ENC20166500.8298VER.1.pdf)

Sessions and Preferences

Methodology / Instrumentation:	**Invited Speaker**
System	**Invited Speaker**
Application / Challenge:	**Invited Speaker**

Options
I am an invited speaker **YES**
I wish to be considered for a talk. **YES**
I am a graduate student. **NO**
I am a graduate student and using this abstract to apply for the Ritchey Award.
Application Form (http://www.enc-conference.org/Portals/0/Ritchey%20Award/Application%20Form%20Ritchey%20Award%20Fill-in%20PDF.pdf?ver=2015-10-28-232642-950) **NO**
I am a postdoc. **NO**

Submitting Correspondent
Bruce Donald

P.O. Box 90129, Department of Computer Science, D101 Levine Science Research Center (LSRC), Research Drive, Duke University

Durham, NC 27708 United States
919-660-6583

brd+enc10@cs.duke.edu

Previous Section

Save Abstract

enc@enc-conference.org, (mailto:enc@enc-conference.org) Phone: (505) 989-4573, Fax: (505) 989-1073 2019
Address: Galisteo Street, Building I-1 Santa Fe, NM 87505

Redefining Structure: New Approaches for (1) Deep Topological Sampling of Protein Structures, and (2) Continuous Probability Density Function for Biomolecular Interdomain Orientations

Yang Qi[1,2], Jeffrey Martin[2], François Thélot[2], Adam Barb[4], Anthony Yan[2], Pei Zhou[1*], Terrence G. Oas[1*], and Bruce R. Donald[1,2,3*]

1 Department of Biochemistry, Duke University Medical Center, Durham, NC.
2 Department of Computer Science, Duke University, Durham, NC; 3 Department of Chemistry, Duke University, Durham, NC.
4 Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA.
* Corresponding Authors

I will describe two novel approaches that redefine the concept of structure from NMR data. The first approach samples fold space using a topological group action on conformation space. It is useful for automated assignment and structure determination of symmetric homoöligomers with ambiguous distance restraints. In the second, we developed a new method for determining the structure of a flexible protein by RDCs that avoids the use of over-parameterized conformational ensembles. Our approach describes the interdomain orientation as a continuous 3D probability distribution, exploiting new geometry, a fitting algorithm, and a new graphical representation (Figure 1, *Right*).

(1) **Ambiguous restraints and Deep Topological Sampling**. Structure determination of symmetric oligomeric complexes presents daunting challenges for NMR structural biologists. In particular, simulated annealing (SA), a widely-used technique, is vulnerable to significant structural errors. Due to assignment ambiguity, SA converges to local minima rather than to the optimal structure or structural ensemble indicated by the data. *Deep Topological Sampling* using *Fold Operator Theory* (DTS/FOT) can overcome these errors, using a systematic search algorithm shown to identify structures that simulated annealing does not find (*Proteins*, 2015). For example, the published NMR and crystal structures of the trimeric membrane protein Diacylglycerol Kinase (DAGK) have very different topologies. Our systematic search techniques not only showed that both published folds are supported by the NMR data, but also found a novel fold that satisfies the data better than either published fold (Figure 1, *Left*). This result highlights the fundamental limitation of global fold determination for homoöligomeric proteins using ambiguous distance constraints. We conclude that the differences in the published NMR and crystal structures are due to limitations of current NMR structure determination methodology. I'll discuss how we overcame these limitations by DTS/FOT to systematically search the space of folds and predict diverse fold topologies that fit all the NMR data well and have good modeling energies.

(2) **Interdomain Dynamics and SpA-N.** Interdomain motions have been modeled previously using discrete conformational ensembles. Instead, we derived a continuous distribution directly from RDCs. We have performed a comprehensive study of the interdomain motions of staphylococcal protein A (SpA), which binds antibodies as part of the pathogenicity of *S. aureus*. We have developed a new NMR-based approach that uses RDCs to determine the interdomain orientational motions of two adjacent SpA domains, and have used our results to explore the relationship between binding energetics and protein dynamics. We believe this is the first time that a continuous joint distribution has been used to represent the interdomain motions of biological macromolecules. This distribution encompasses the continuous dynamic nature of these macromolecules and redefines structure in terms of a probability function of interdomain orientations (Figure 1, *Right*). Using this new paradigm, we capture the smooth transitions between important states and deduce functionally relevant properties that previous views have overlooked. This continuous distribution allows us to straightforwardly calculate the re-orientational components of the free energy, internal energy and entropy of antibody binding. We conclude that the observed distribution of interdomain orientation is biased in favor of the probable bound orientation. Our study represents a first step toward a quantitative estimate of the role of conformational rearrangements in macromolecular interactions.

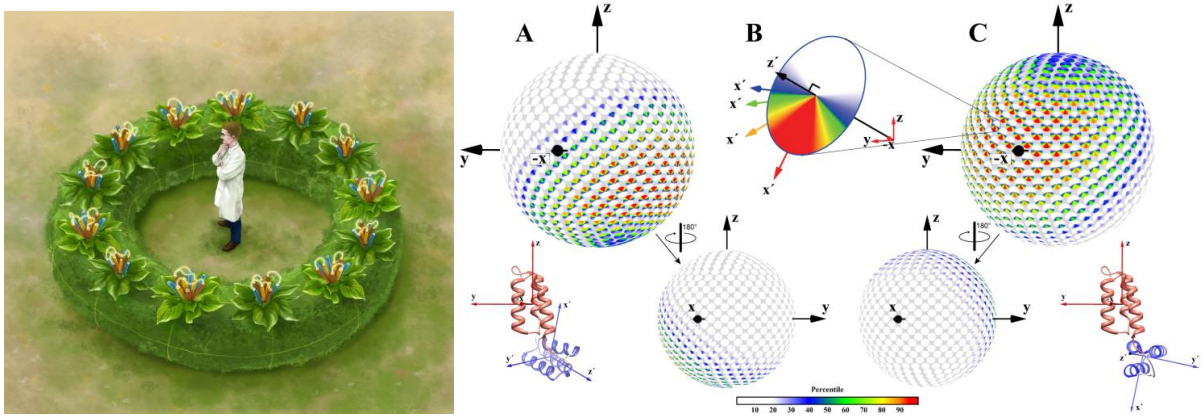

Figure 1. *Left:* Concept of a large number of different folds of DAGK, which all plausibly satisfy the NMR data. The toroidal configuration space illustrates how Fold-Operator Theory obtains a systematic search algorithm over the different possible folds. We report 7 new folds that are topologically distinct, fit the data well, and differ by up to 12 Å (transmembrane helix backbone RMSD) from either the previously published NMR structure or the crystal structure. Image: *Proteins* 2015; 83(4): 651-661 (Cover). *Right:* Continuous distribution of interdomain orientations (CDIO) models for a di-domain construct of SpA-N shown in disk-on-sphere (DoS) views. Two solutions (panels A and C) give equivalently good fits to the SpA-N data. Also shown are atomistic models whose interdomain orientation is the most probable in each solution. (B) An example disk showing the joint probabilities of 4 different interdomain orientations, all with the same *z'* axis orientation but different rotations around *z'*.