28 02-Methods for Structure Determination and Analysis, Computing and Graphics

MS-02.01.03 PROGRESS TOWARDS APPLICATION OF THE MINIMAL FUNCTION TO MACROMOLECULES. Charles M. Weeks*[1], Russ Miller[2], G. David Smith[1], Martha M. Teeter[3] & Herbert A. Hauptman[1]. [1]Medical Foundation of Buffalo, 73 High St., Buffalo, NY 14203, USA; [2]Dept. of Computer Science, State University of New York at Buffalo, Buffalo, NY 14260, USA; [3]Dept. of Chemistry, Boston College, Chestnut Hill, MA 02167, USA.

The *Shake-and-Bake* method of structure determination consists of a phase refinement procedure based on Hauptman's minimal function, R(φ), alternated with Fourier filtering [Weeks, C.M. *et. al.* (1993). *Acta Cryst.* **D49,** 179-181; Miller, R. *et. al.* (1993). *Science*, in press.]. This method provides a powerful and convenient formulation of direct methods, having been used to solve several known and unknown peptide structures containing approximately 100 non-hydrogen atoms in the asymmetric unit. It has also been applied successfully to atomic resolution data for two previously known small proteins, gramicidin A [Langs, D.A. (1988). *Science*, **241**, 188-191.] and crambin [Hendrickson, W.A. & Teeter, M.M. (1981). *Nature*, **290**, 107-113.]. These structures contain approximately 300 and 400 atoms, respectively, after taking disorder and partial occupancy into account.

In the case of gramicidin A, a 0.12Å grid placed in the asymmetric portion of the unit cell was used to obtain initial coordinates for 240,396 1-atom trial structures. The 346 trials having an initial mean phase error ≤ 80° were subjected to the *Shake-and-Bake* procedure, and three solutions were obtained following 450 cycles of refinement and filtering. Thus, in the worst case scenario, there is one solution per 80,000 trials for gramicidin A. In the case of crambin, initial phases were obtained by performing structure factor calculations based on

	Gramicidin A	Crambin
Resolution	0.86Å	0.83Å
Temperature	120°K	130°K
Space Group	$P2_12_12_1$	$P2_1$
Atoms (approximate)	300	400
Phases	2000	4000
Triplets	20,000	40,000
Negative Quartets	0	0
Cycles	450	200
Trials Generated	240,396	1216
Trials Processed	346	1216
Solutions	3	16

randomly positioned 2-atom trial structures. The success rate was 1.3% following 200 *Shake-and-Bake* cycles. Both the gramicidin A and crambin maps can be easily interpreted either by examination of interpeak distances and angles or by graphical electron density fitting using FRODO. The best maps for manual examination are obtained by terminating the procedure with one cycle of Fourier refinement using all statistically reliable measured data.

At present, these experiments leave several questions unanswered. For example: What is the random start success rate for gramicidin A? How important was the presence of the six sulphurs in the crambin application? How long will the procedure, which presently relies on peak picking at the Fourier stage, be applicable as the resolution of the data is decreased? These problems are presently under investigation.

This research was supported by NIH grant GM-46733 and by NSF IRI-9108288. The authors also thank Dr. R. H. Blessing for help with data reduction and Dr. R. Jones of the Thinking Machines Corporation for computer time.

MS-02.01.04 **MAXIMUM ENTROPY, LIKELIHOOD AND THE CRYSTALLOGRAPHY OF BIOLOGICAL MACROMOLECULES** By C.J.Gilmore', G.Proctor, and J.R.Fryer *Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland*, G.Bricogne, *LURE, Bâtiment 209D, Orsay 91405, France*, S.Xiang and C,W.Carter Jnr. *Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, North Carolina, USA*, and A.Brisson, G. Moser and M. Schmutz, *Inistut de Chimie Biologique, Strasbourg, France.*

The maximum entropy-likelihood method as formulated by Bricogne (Bricogne, *Acta Cryst.* (1984) **A40,** 410-445) and implemented by Bricogne and Gilmore (*Acta Cryst.* (1990) **A46,** 284-297), and Xiang, Carter, Bricogne and Gilmore (*Acta Cryst.* (1990) **D47,** 193-212), provides a powerful phasing tool for use in macromolecular crystallography including electron crystallography. We report and summarize here the following:
(1) The use of likelihood to select correct phase sets generated by the SAYTAN program for the protein avian pancreatic polypeptide. App. (Gilmore, C.J., Henderson, A.N. & Bricogne, G. *Acta Cryst.* (1991) **A47,** 842-846),.
(2) The use of entropy maximisation combined with solvent flattening applied to cytidine deaminase.(Xiang, S., Carter, C.W., Bricogne, G. and Gilmore C.J.(*Acta Cryst.* (1990) **D47,** 193-212).
(3) Two structures of biological macromolecules studied using electron diffraction, and phased image data from high resolution electron microscopy:

(a) Purple membrane (*Halobium Halobactorium*) data (Baldwin, J.M., Henderson,R., Beckman, E., & Zemlin, F. *J. Mol. Biol.* (1988) **202,** 585-591). This was as a test of the method, and has produced some controversial results concerning resolution enhancement.
(b) Cholera toxin. Here we are phasing data to 4Å from 56 unique phased reflections at 8.8Å resolution using the ME method incorporating the application of five-fold non-crystallographic symmetry, and solvent flattening.

In both (a) and (b) a low resolution basis set of phased reflections had been derived from the Fourier transform of optical image data suitably averaged, and used to phase the high resolution diffraction data *via* a process of entropy maximisation and likelihood evaluation coupled with the building of phasing trees.

The maximum entropy method is ideal in these circumstances because:
(1) It will work with projection data.
(2) It is stable regardless of data resolution.
(3) It can utilize non-crystallographic symmetry, and solvent flattening in a wholly natural and relatively simple way.
(4) It uses non-uniform atomic distributions which are constantly updated.
(5) Likelihood can be used to determine an effective unit cell contents that reflects the data resolution.

MS-02.01.05 ELECTRON-DENSITY HISTOGRAMS AND THE PHASE PROBLEM. By V Yu Lunin, Institute of Mathematical Problems of Biology, Pushchino, Moscow Region, Russia.

The spectra of frequencies (histograms) of different values in Fourier syntheses provide the most adequate representation of information on 'what values may be found in a good Fourier synthesis and how frequent they are'. These Electron-Density Histograms (EDH)

