MS02.06.04 AB INITIO PHASE DETERMINATION FOR VIRUSES; THE USE OF NON-CRYSTALLOGRAPHIC SYMMETRY FOR PHASE REFINEMENT. Chapman, M.†‡, Tsao, J.*‡, Rossmann, M.‡, Munshi, S.¶‡, Johnson, J.§‡, †Dept. Chem. & Inst. Molec. Biophys., Florida State Univ., Tallahassee, FL 32306; *Dept. Microbiol., Univ. Alabama, Birmingham, AL 35394; ‡Dept. Biol. Sci., Purdue Univ., W. Lafayette, IN 47907; ¶MRL, WP144B-122; Merck & Co. Inc., W. Point, PA 19486; §Dept. Molec. Biol., Scripps Res. Inst., La Jolla, CA 92037

The atomic structure of Canine Parvovirus (CPV) was determined with the use of phases calculated *ab initio* [Tsao *et al.*, 1992, *Acta Cryst.*. **B48**:75]. *Ab initio* phases were refined and extended to 8 Å and used to solve the positions of a heavy atoms for a single isomorphous replacement (SIR) structure solution. In retrospect, it was shown [Tsao *et al.*, 1992, *Acta Cryst.* **A48**:293] that the phases could be extended from 8 to 3.25 Å and that the virus structure could have been determined *ab initio* without SIR. Similar approaches have since been used to determine the structures of ϕX174 [McKenna, *et al.*, 1992, *Acta Cryst.* **B48**:499] and Nωv.

With the advantage of phase refinement and extension using non-crystallographic symmetry of high redundancy (60-fold for CPV), crude centric approximations can suffice for the initial starting phases at low resolution. Approximating the virus by a spherical shell of uniform densities for protein and nucleic acid, the sign of the molecular transform oscillates with the positions of nodes highly dependent on the choice of shell radii. *Post mortem* examination of successful and failed attempts show that it is critical to have a self consistent set of phases in the resolution shell from which extension will take place. This means that the radii must be estimated to within ~1% of their true values (or be similarly consistent with an alternative Babinet solution). Although trial-and-error may eventually give a suitable initial phasing model, it is also possible to refine the radii and position of the spherical shell model through optimization against low resolution amplitudes from crsytallography or solution scattering [Chapman et al., 1992, *Acta Cryst.* **A48**:301].

MS02.06.05 GENETIC ALGORITHMS AND MACROMOLECULAR PHASING M. Fujinaga & M. N. G. James MRC of Canada, Group in Protein Structure and Function Department of Biochemistry, University of Alberta Edmonton Alberta, Canada T6G 2H7

A program has been developed for exploring the use of genetic algorithms for macromolecular phasing. Genetic algorithms are powerful optimization techniques that borrow ideas from natural evolution. Unlike normal optimization techniques, it deals with a set (population) of possible solutions and these are improved by combining (mating) pairs of solutions from the population. The choice of the pair of solutions to mate is governed by the function value that one is trying to optimize (fitness function). The implementation of a genetic algorithm for the crystallographic problem has been done by representing a set of carbon atoms on a three-dimensional grid. The fitness function includes contributions from the agreement between observed and calculated structure factors as well as conformity to the expected distribution of atoms in a protein. One of the advantages of genetic algorithm is that it does not rely on gradients so that functions without derivatives can be included in the fitness function. The mating is done by a continuous crossover method that involves complete mixing and random separation of two sets of atoms. The number of atoms in common is monitored and the offspring are penalized according to the degree of inbreeding. The entire process is rather computationally expensive and the program has been parallelized to run on 20 IBM RS6000 workstations linked together using PVM (parallel virtual machine). The method is being tested on the complex of TEM-1 beta-lactamase and beta-lactamase inhibitory protein (BLIP). The enzyme part of the structure (62%) had been solved by molecular replacement but it was not possible the locate the inhibitor. Density modification techniques have failed to improve the phases sufficiently to solve the remaining 38% of the structure. The structure was eventually resolved by molecular replacement and refined so that a set of 'correct' phases exists to monitor the progress of the genetic algorithm.

This work is supported by the Medical Research Council of Canada.

MS02.06.06 PHASING OF PEPTIDE AND PROTEIN DATA SETS BY AB INITIO MODELLING. Piet Gros, Dept. of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 18, 3584 CH Utrecht, The Netherlands (gros@chem.ruu.nl)

A method is presented for ab initio phasing of protein and peptide data sets at medium resolution; examples of phasing peptide data sets up to low resolution (ca. 4.5 Ang) will be presented. The optimization method consists primarily of solvent flattening and structure refinement (using Molecular Dynamics and Energy Minimization). To make structure refinement feasible for random starting models a force field is introduced that is applicable to loose and all equal atoms. Pseudo-energy potential functions (potentials of mean force) in this force field are derived from radial-distribution functions of all non-hydrogen atoms as is observed in refined structures. Thus, the configuration of atoms is restrained towards atomic distributions of protein structures. Solvent flattening is applied to estimate the bulk solvent contribution to the structure factors. This procedure has been tested for a few peptide data sets. The data was generously supplied by Dr Isabella Karle. In the tests the data were truncated to 2.5 Ang resolution, because data to atomic resolution is rarely available when a protein structure is being determined. The resulting models from these optimization show a correspondence at low resolution to the known structures. Analysis of the phase differences and map correlation coefficients (w.r.t. the known answers) indicates that phase information up to approx. 4.5 Ang is obtained.

MS02.06.07 THE *AB INITIO* STRUCTURE AND REFINEMENT OF A SCORPION PROTEIN TOXIN. G.D. Smith[1,2], R.H. Blessing[1], S.E. Ealick[3], J.C. FontecillaCamps[4], H.A. Hauptman[1], D. Housset[4], D.A. Langs[1], R. Miller[1,5]. [1]Hauptman-Woodward Med. Res. Inst., 73 High St., Buffalo, NY 14203 USA; [2]Roswell Park Cancer Institute, Elm & Carlton St., Buffalo, NY 14263 USA; [3]Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853 USA; [4]Laboratoire de Cristallographie et Cristallisation des Proteines, Institut de Biologie Structurale Jean Pierre Ebel-CEA, CNRS, 41, Avenue des Martyrs, 38027 Grenoble Cedex 1 France; [5]Dept. of Computer Science, State Univ. of New York at Buffalo, Buffalo, NY, 14260 USA.

The structure of a 64 residue scorpion toxin, which crystallizes in space group $P2_12_12_1$, has been determined *ab initio* at 0.96Å using the *SnB* program. A total of 50,000 triples were generated from 5,000 phases and the value of R_{min} clearly revealed a single solution amongst the 1619 trials, using 255 *Shake-and-Bake* cycles per randomly generated trial structure. The number of residues in the protein and the amino acid sequence were deliberately withheld from the Buffalo group; the only known information was that the protein was composed of approximately 500 atoms and contained 4 disulfide bonds. A very conservative initial starting model consisted of five fragments varying in length from 5 to 18 residues for a total of 241 atoms. The entire structure was revealed following multiple cycles of Xplor refinement and Fourier