# Refinement of Herpesvirus B-Capsid Structure on Parallel Supercomputers

Z. H. Zhou,* W. Chiu,# K. Haskell,* H. Spears, Jr.,§ J. Jakana,# F. J. Rixon,¶ and L. R. Scott*
*Texas Center for Advanced Molecular Computation, University of Houston, Houston, Texas 77204-3476 USA; #Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030 USA; §Silicon Graphics, Houston, Texas 77077 USA; and ¶MRC Virology Unit, Institute of Virology, Glasgow G115JR, Scotland

ABSTRACT Electron cryomicroscopy and icosahedral reconstruction are used to obtain the three-dimensional structure of the 1250-Å-diameter herpesvirus B-capsid. The centers and orientations of particles in focal pairs of 400-kV, spot-scan micrographs are determined and iteratively refined by common-lines-based local and global refinement procedures. We describe the rationale behind choosing shared-memory multiprocessor computers for executing the global refinement, which is the most computationally intensive step in the reconstruction procedure. This refinement has been implemented on three different shared-memory supercomputers. The speedup and efficiency are evaluated by using test data sets with different numbers of particles and processors. Using this parallel refinement program, we refine the herpesvirus B-capsid from 355-particle images to 13-Å resolution. The map shows new structural features and interactions of the protein subunits in the three distinct morphological units: penton, hexon, and triplex of this $T = 16$ icosahedral particle.

## INTRODUCTION

Caspar and Klug's original theory of quasi-equivalence (Caspar and Klug, 1962) provided an enormous intellectual stimulus to understanding the principles of virus assembly (Casjens, 1997) and to the development of biophysical techniques for structural determination of viral assemblies (Baker and Johnson, 1997). Determination of the high-resolution crystallographic structures of plant and animal icosahedral viruses has shown that the structural variation among the quasi-equivalent subunits often differs significantly from that predicted by the Caspar and Klug theory (Harrison et al., 1996; Johnson, 1996).

Whereas x-ray crystallography has long been a powerful technique for revealing atomic-scale structural details of viral capsid proteins, in the past few years electron cryomicroscopy has played a major role in moderate-resolution structure determination for viruses with diameter ranges of 250-1300 Å. These studies have yielded new insights into a number of fundamental structural aspects of viruses, including the assembly mechanisms (e.g., Conway et al., 1995; Dryden et al., 1993; Prasad et al., 1993; Stewart et al., 1993), the roles of scaffolding proteins (e.g., Marvik et al., 1995; Thuman-Commike et al., 1997), the interactions between ordered nucleic acids and coat proteins (e.g., Cheng et al., 1994; Prasad et al., 1996), the sites of cellular receptor binding (e.g., Olson et al., 1993), and the mechanisms of antibody neutralization (e.g., Chiu and Smith, 1994; Smith et al., 1995). Recent technical advances in electron cryomicroscopy and image processing have made it possible to determine the three-dimensional (3D) structures of moderate-sized icosahedral particles to resolutions at and beyond 9 Å, at which point $\alpha$-helices within the proteins can be seen (Böttcher et al., 1997; Conway et al., 1997; Trus et al., 1997).

The herpes simplex virus-1 (HSV1) capsid is among the largest and most complex icosahedral particles. Its structural determination to high resolution will require improvements in both image acquisition and data-processing techniques. This paper addresses our approach to improving the resolution of the HSV1 B-capsid structure. In the next section we describe the overall experimental and computational approaches we have employed and identify global refinement of particle orientations as a major limiting step that is suitable for parallel computation. In the third section we discuss the factors influencing the potential performance of parallel computation procedures, and describe our implementation on different shared memory supercomputers and the parallel performance obtained. The fourth section describes a 13-Å structure of the HSV1 B-capsid, obtained by using the parallel refinement. In the fifth section we discuss the advantages of our experimental and computational approach to solving higher resolution structures of large icosahedral particles, with particular reference to the use of parallel computing.

*Received for publication 18 April 1997 and in final form 7 October 1997.*

Address reprint requests to Dr. Wah Chiu, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. Tel.: 713-798-6985; Fax: 713-796-9438; E-mail: wah@bcm.tmc.edu.

Dr. Zhou's present address is University of Texas–Houston Medical School, Department of Pathology and Laboratory Medicine, Houston, TX 77030.



0006-3495/98/01/576/13 $2.00

## THREE-DIMENSIONAL RECONSTRUCTION OF ICOSAHEDRAL PARTICLES

The 3D reconstruction method for electron microscopy was pioneered by DeRosier and Klug (1968). Their method combines projection images of different views of an object, where the number of required views ($N$) depends on the size of the objects ($D$) and the targeted resolution ($d$), i.e., $N \geq$

$\pi D/d$. In the case of an icosahedron, one needs fewer views than for an irregular object because of the redundancy contributed by the 60 identical asymmetrical units that make up each particle (Crowther, 1971; Crowther et al., 1970). For instance, to reconstruct a 1000-Å-diameter icosahedral particle at 5-Å resolution, only 10 evenly sampled, noise-free views are, in theory, required. In practice, images are noisy and have low contrast, so that hundreds, even thousands of such images are needed to obtain statistically well-defined structural details.

The algorithm for 3D reconstruction of icosahedral particles embedded in negative stain was first introduced by Crowther and co-workers (Crowther, 1971; Crowther et al., 1970). Since then, there have been improvements both in the data-processing procedure and in Crowther's original code (e.g., Baker et al., 1988; Fuller, 1987; Prasad et al., 1988), to permit processing of low-contrast images of ice-embedded particles. To improve the computational efficiency and allow for the increased complexity of larger particles at higher resolution, several computational methods have been developed for the selection of particles (Thuman-Commike and Chiu, 1996), the estimation (Olson and Baker, 1989; Thuman-Commike and Chiu, 1997) and refinement (Baker and Cheng, 1996; Fuller et al., 1996) of particle centers and orientations, and the correction of the contrast transfer function (CTF) of the microscope (Böttcher et al., 1997; Conway et al., 1997; Trus et al., 1997).

One of the most demanding tasks in image processing is to obtain accurate values for the two center and three orientation parameters of each particle image. The accuracy required depends on the desired resolution and the particle size. For example, the maximum allowable error in the center and orientation determination must be less than ~2.5 Å and ~0.3°, respectively, to produce a 10-Å map of a 1000-Å-diameter particle. This estimate is based on the criterion that the maximum amount of phase shift in the computed Fourier transform of the particle image introduced by a change of center and orientation parameters of these amounts at the targeted resolution should be less than 90°.

## Experimental and computational strategy

In our experimental set-up, a JEOL4000 intermediate voltage electron cryomicroscope is operated at 400 kV using a $LaB_6$ filament, and the ice-embedded specimen is kept at −170°C. The intermediate-voltage electron microscope offers the prospect of reduced chromatic effects (Brink and Chiu, 1991) and a greater depth of field (Chiu et al., 1986; Zhou and Chiu, 1993). Our images are recorded by a spot-scan procedure to minimize beam-induced movement (Downing, 1991; Zhou et al., 1994). Our experimental data consist of focal pairs comprising close-to- and out-of-focus

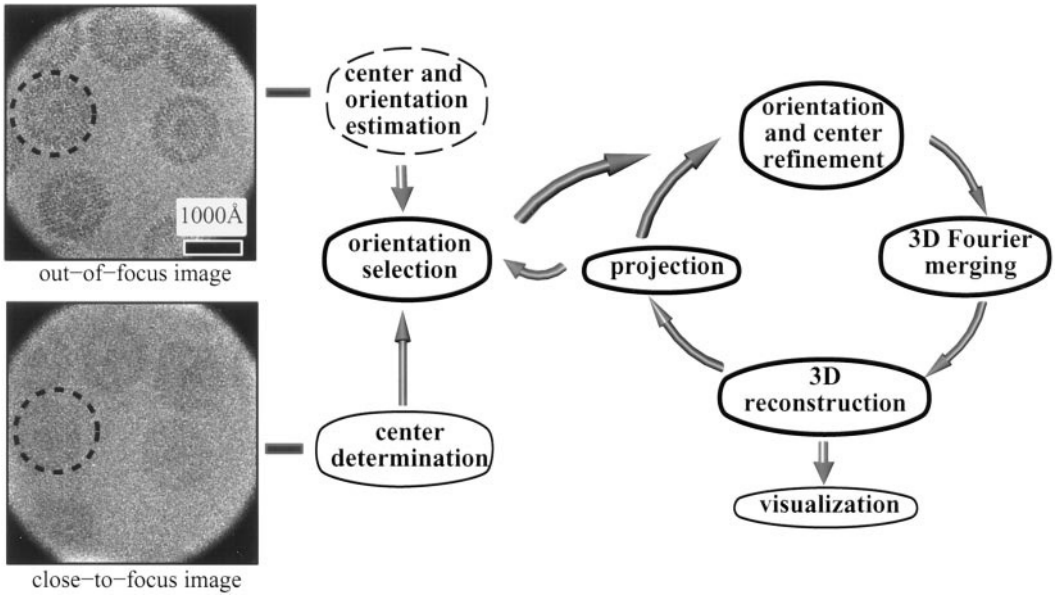


FIGURE 1 Flow-chart diagram illustrating the data-processing scheme for 3D structural determination of icosahedral virus particles from spot-scan images. On the left are portions of a focal pair of 400-kV, spot-scan micrographs of HSV1 B-capsid particles with one corresponding capsid particle image indicated by dashed circles in both images. Data-processing steps are listed on the right, inside ovals of dashed (for out-of-focus particles), thin solid (for close-to-focus particles), and thick solid (for both out-of-focus and close-to-focus particles) lines. The procedure begins with out-of-focus particles, first obtaining a set of possible orientation estimates for each particle by self common-line searches, followed by elimination of incorrect orientation estimates, then orientation and center refinement, 3D Fourier merging, 3D reconstruction, and computation of 2D projections. The center of each close-to-focus particle image is first determined by cross-correlating the particle image with a computed projection, using the orientation estimated from the corresponding out-of-focus particle. Iterative cycles of orientation and center refinement, merging, and 3D reconstruction are then carried out by using the close-to-focus particle images. The refinement steps include both local (with the raw images against a reprojection template) and global (with only raw images) refinement. The final step is 3D visualization.

images (Fig. 1). The former contains high-resolution but low-contrast image data for use in the eventual high-resolution reconstruction, whereas the latter contains particle images of relatively high contrast from which a preliminary low-resolution 3D reconstruction can be readily obtained (Zhou et al., 1995). Computed projections of this initial low-resolution 3D reconstruction are reasonably noise-free and are used to help by determining the center and orientation parameters of the corresponding close-to-focus particle images.

We have used iterative algorithms to refine the center and orientation parameters (Fig. 1). The first refinement method is local in that particle parameters are refined against templates consisting of a set of particle images or computed projections. At the initial stage of data processing, the templates are raw particle images for which parameters have been estimated. When a preliminary 3D map is available, its computed projections would be used as templates for the local refinement. However, because there is no a priori guarantee that the local refinement alone will achieve the desired accuracy, the iterative refinement shown in our data-processing scheme (Fig. 1) includes both local and global refinement steps. In the global refinement step, all of the raw particle images are refined against each other, without using templates. A potential merit of global refinement is the absence of possible bias arising from template data. However, it is a computationally intensive process, particularly with a large data set.

## Determination of particle parameters by the use of common-lines

The principle of 3D reconstruction is based on the central projection theorem illustrated for a model system in Fig. 2, which states that the Fourier transform of a two-dimensional (2D) projection of a 3D object is equivalent to a central section through the 3D Fourier transform normal to the direction of the projection (Fig. 2 *c*). Therefore, two Fourier sections of images of particles with different orientations must intersect. The two lines coinciding at this intersection, one in each transform, are called a pair of *cross common-lines* (Fig. 2 *d*). Along the cross common-lines, the amplitudes and phases of the Fourier coefficients at corresponding points, one in each transform, are equal for noise-free images (Crowther, 1971; Crowther et al., 1970).

Each icosahedron consists of 60 asymmetrical units related by five-, three-, and twofold symmetry. Correspondingly, there are 60 Fourier central sections per icosahedron that are symmetry-related (Fig. 2 *c*), each of which intersects the Fourier section of the other particle image. Thus there are a total of 60 pairs of cross common-lines between any two particle transforms (Fig. 2 *d*). However, when the incident electron beam coincides with one of the symmetry axes, this number decreases because of degeneracy. For any one particle image, the symmetry-related sections also intersect with the original section, resulting in common-lines

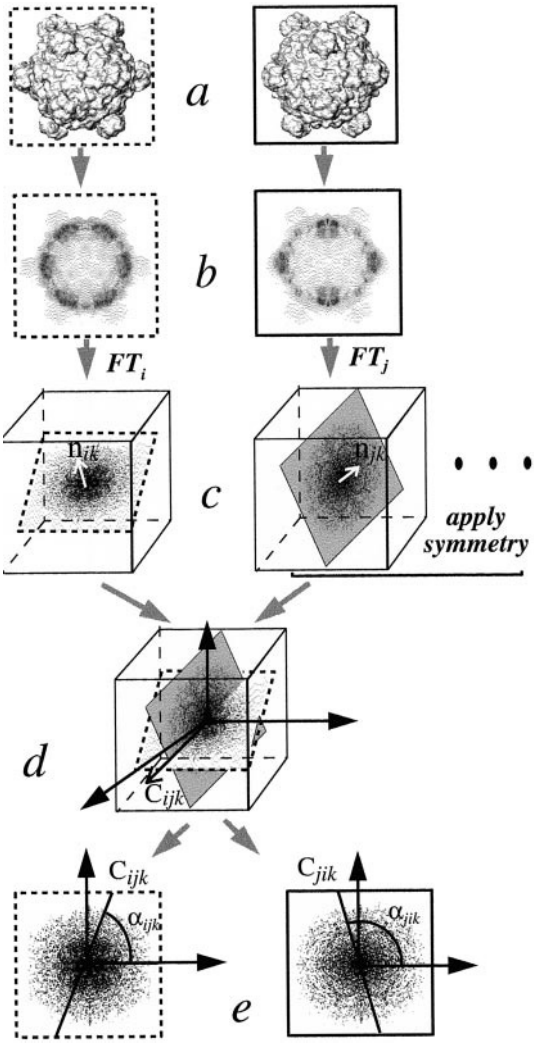


FIGURE 2 Illustration of cross common-lines. (*a*) An icosahedral model is generated from the atomic coordinates of one asymmetrical unit of bacterial phage $\phi$X174 (McKenna et al., 1992) obtained from the Brookhaven Protein Data Bank. (*b*) Projection images of the model are generated for two different viewing directions. (*c*) Each 2D Fourier transform *i* or *j* is equivalent to a central section of the Fourier transform of the 3D model. The vectors normal to these sections, $\mathbf{n}_{ik}$ and $\mathbf{n}_{jk}$, are parallel to the projecting directions. (*d*) These sections intersect at one line represented by a directional unit vector $\mathbf{C}_{ijk}$. The Fourier values along the intersection at corresponding spatial frequency *s* in each section are identical. (*e*) These lines, which are located at angular positions $\alpha_{ijk}$ and $\alpha_{jik}$ in the original sections, are called a pair of cross common-lines. Icosahedral symmetry allows the Fourier section *j* in *c* to be rotated to generate another 59 sections, each of which also intersects with Fourier section *i*, resulting in a total of 60 pairs ($k = 1 \rightarrow 60$) of cross common-lines between transforms *i* and *j*. For simplicity, the symmetry-related sections of section *j* are not shown in *d*, and only one of the 60 pairs of common-lines is drawn in *e*.

that are called *self common-lines*. Because of degeneracy, there are only 37 or fewer unique pairs of self common-lines within each transform (Fuller et al., 1996).

The directional parameter of a self or cross common-line can be mathematically described as a unit vector, $\mathbf{C}_{ijk}$, for particle *i*, which can be expressed as the cross-product of the

two directional vectors, $\mathbf{n}_{ik}$ and $\mathbf{n}_{jk}$, which are normal to the two central sections (Fig. 2 *c*). The subscripts in these parameters refer to particles *i* and *j* and their common-line *k*. In the case of a self common-line, $i = j$, and *k* refers to the self common-line generated from the symmetry-related sections of the particle. In the case of a cross common-line, $i \neq j$, and *k* refers to the cross common-line between the sections of particle *i* and the *k*th symmetry-related sections of particle *j*. Similarly, $\mathbf{C}_{jik}$ is defined as the directional parameter for the common-line *k* for the intersections of the section of particle *j* with those of the particle *i*, and coincides with $\mathbf{C}_{ijk}$ in Fig. 2 *d*. The distributions of the self common-lines are dependent on the projection direction of an individual particle, whereas those of the cross common-lines depend on the projection directions of two particles. In principle, the locations of 37 or fewer self common-lines alone are sufficient for determining the orientation of an icosahedral particle. However, using all cross common-lines from a large set of particle images provides a larger data set to impose a self-consistent constraint for their orientation determination and refinement. The major computational effort in the determination of the particle orientation parameter is devoted to determining the $\mathbf{C}_{ijk}$, and thus $\mathbf{C}_{jik}$, as accurately as possible by minimizing the phase differences between the data points along all pairs of the common-lines in the Fourier transforms from a large set of noisy particle images.

The algorithm for determining the orientation ($\phi$, $\theta$, $\omega$) and center ($x$, $y$) parameters proceeds as follows. For each particle, the phase residuals along the common-lines corresponding to each orientation in a set of trial orientations are evaluated. For initial orientation estimation, the trial orientations are sampled on a grid whose size is typically 1° covering all possible orientations. Because of the 60 copies of the asymmetrical unit in an icosahedral particle, the unique possible ranges of the angles are $\phi = -31.72° \rightarrow 31.72°$, $\theta = 69.09° \rightarrow 90°$, and $\omega = 0° \rightarrow 360°$. For refinement steps, the trial orientations are sampled at smaller grid sizes, and the range of the angles is restricted to within 1–2° around the initial orientation estimate. Because a change in the particle center results in a phase shift in the Fourier data along the common-lines, the center parameters of each particle must also be determined accurately and are refined simultaneously with the angles.

In our software, the average phase residual $R_i$ along the self and cross common-lines between particle *i* and all other particles is evaluated according to Eq. 1. The angles ($\phi$, $\theta$, $\omega$) and center ($x$, $y$) that minimize this function will be taken as the correct orientation and center parameters for particle *i*.

$$R_i(\phi, \theta, \omega, x, y) = \frac{\sum_{j=1}^{N} \sum_{k=1}^{k_{max}(j)} \sum_{s=s_{min}}^{s_{max}} |\psi_i(s, x_i, y_i, \alpha_{ijk}) - \psi_j(s, x_j, y_j, \alpha_{jik})| \times \delta(s, \alpha_{ijk}, \alpha_{jik})}{(s_{max} - s_{min}) \sum_{j=1}^{N} k_{max}(j)} \quad (1)$$

where $\psi_i(s, x_i, y_i, \alpha_{ijk})$ is the phase value in the Fourier transform of the *i*th particle along the *k*th common-line; $\psi_j(s, x_j, y_j, \alpha_{jik})$ is the phase value in the Fourier transform of the *j*th particle (or of its symmetry-related sections) along the *k*th common-line; $\alpha_{ijk}$ and $\alpha_{jik}$ are the angles of the *k*th common-lines in the particle's Fourier transforms; $s_{min}$ and $s_{max}$ define the frequency range within which the phase residuals are evaluated; $k_{max}$ is the total number of common-lines (either 37 for self common-lines or 60 for cross common-lines when there is no degeneracy); *N* is the number of particles; $\delta$ defines a weighting function for the amplitude threshold, which could be either 1 or 0, depending on the Fourier amplitude at different frequencies.

## Computation time and memory requirement for global refinement

Global refinement of the initial estimates of orientation and center parameters for each particle *i* is done by minimizing $R_i$ (as defined in Eq. 1) along the self and cross common-lines between particle *i* and all other (*j*) particles, using a coordinate descent method (Brent, 1973). Our program emicograd_org carries out the global refinement and is a modification of the original orientation refinement program *emicograd* (Fuller, 1987), which has been enhanced to incorporate simultaneous center refinement (Zhou, 1995). The computation time required for global refinement limits the practicality of carrying out reconstructions of large particles, because it increases quadratically with the number of particles, whereas the computation time for initial orientation search and local refinement grows linearly. For example, a global refinement of 200 particles 1250 Å in diameter takes ~2 days on an SGI R4400 workstation. By extrapolation, a global refinement of 1000 particles would take more than 50 days.

Another constraint on the use of the global refinement calculation is the memory requirement. There are two kinds of memory allocation in *emicograd_org*: static memory, which depends only on the maximum number of particles determined at the time when the program is compiled in a computer; and dynamic memory, which depends on the size of the particle images, the number of particles actually used, data sampling parameters, and the desired resolution of the refinement. The bulk of the memory is used to store the Fourier transforms of particle images and the common-line phase residual matrices, all of which must be accessible at run time to perform the minimization. Based on the analysis of our code, the number of computer words used by the static memory allocation is

$$M_s = 3N_{max}^2 + 77N_{max} + C \quad (2)$$

where $N_{max}$ is the maximum number of particle images that could be solved by the program, and *C* is a constant on the order of 100 words. The number of words dynamically allocated is

$$M_d = 4N^2S + 2NS^2 + 8N^2 + 126NS + 244N \quad (3)$$

where $N$ is the actual number of particles being solved and $S$ is the maximum spatial frequency in pixel units. For example, assuming each word is 4 bytes, the number of particles used is 355, and the resolution is 105 pixels (equivalent to 1/13 $Å^{-1}$ for HSV1 B-capsid), the dynamic memory allocation is 266 Mbytes. With a further increase of $N$ to 800 and $S$ to 150-pixel resolution (necessary to achieve ~9-Å resolution for the HSV1 B-capsid), the dynamic memory allocation requirement will become 1.76 Gbytes. This amount of RAM is not typically available on a single processor, but is commonly available on multiprocessors. The use of disk storage for such computation is a possible alternative, but execution time would be substantially longer.

## REFINEMENT BY PARALLEL PROCESSING

Based on the above discussion, it is apparent that the execution of global refinement with hundreds to thousands of icosahedral particles requires high-performance computers with multiple processors and large memories. In this section, we will review the computer hardware we have considered for our implementation, discuss performance issues for parallel applications, and then describe our implementation of the parallel global refinement algorithm. Finally, we present the performance results achieved.

### Parallel computing overview

There are two principal types of parallel computer hardware architecture: a distributed memory multicomputer and a shared-memory multiprocessor (Lewis and El-Rewini, 1992; Wilson, 1995). A distributed memory multicomputer consists of multiple processors, each with its own memory, connected via any of a variety of network topology choices (Fig. 3 *a*). A shared memory multiprocessor consists of multiple processors, each usually having local cache memory, connected to a shared common memory via a bus or switch (Fig. 3 *b*).

There are two categories of memory model for parallel programming, also called *shared* and *distributed*, which differ conceptually in ways that are analogous to the hardware architecture models. However, there is theoretically no restriction against compiling a distributed-memory program on a shared-memory computer or vice versa. The shared-memory model implies uniform access to all program memory by all processors, whereas the distributed-memory model implies no direct access to a processor's memory by other processors. In Fig. 3 *c* we illustrate the combination of a shared-memory programming model implemented on a distributed-memory multicomputer, which is often referred to as a *virtual shared memory* or *distributed shared memory* computer. One practical advantage of using a shared-memory programming model is ease of programming, because it does not require the user to specify explicitly the distribution of the data among processors.

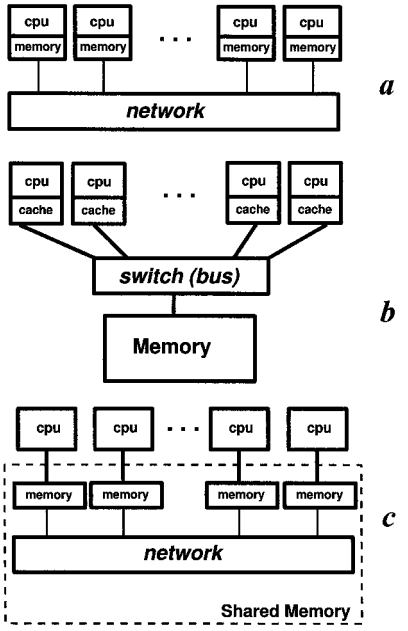


FIGURE 3 Schematic models of computer hardware architecture. (*a*) Distributed memory multicomputer. (*b*) Shared-memory multiprocessor. (*c*) Shared-memory model implemented on distributed memory multicomputer.

### Factors affecting parallel performance

Speedup and efficiency are commonly used measures of parallel performance. The *speedup* ($S_p$) is defined as the ratio of the single processor time to the elapsed time (wall clock time) for solving the same problem on $P$ processors. Whereas speedup measures how much faster a parallel program runs on a particular computer, it does not measure the effective use of the processors. The efficiency of a program on $P$ processors is the ratio, $S_p/P$, of the speedup achieved to the number of processors used to achieve it (Wilson, 1995).

The first steps in parallelizing an application are to identify the most time-consuming parts of the application, and to minimize the amount of work that must be done sequentially. However, there are other factors, such as load balance and granularity, that affect the potential speedup and efficiency. The load balance is defined as the average computing time on all processors, divided by the maximum time by any processor. Granularity refers to the amount of computation distributed to processors. As a general rule, greater efficiency is achieved with coarse-grained parallelism by distributing large blocks of work to different processors, provided that a good load balance can be maintained (Wilson, 1995). In a real application, it is rarely feasible to attain perfect load balance (i.e., an even distribution of computing time on all processors). Different job-scheduling protocols can be used to optimize the load balance for specific applications.

There are two major scheduling strategies for distributing work across processors. The static scheduling strategy determines how the work will be distributed among processors

at the beginning of the execution of the program, and is the default scheduling method on many parallel computers. It is appropriate when the pieces of work to be distributed are of equal size and thus presumably require the same amount of time to complete. In the dynamic scheduling strategy, work is distributed to the processors on a first-come, first-served basis at run time. This strategy typically results in better load balancing among processors when the pieces of work to be distributed do not require the same execution times (Lewis and El-Rewini, 1992; Wilson, 1995). For example, if there are two jobs that each take a large amount of time and many tasks that each take a small amount of time, the dynamic scheduling strategy ensures that the two large jobs will not be executed by the same processor.

### Parallel implementation of global refinement

Using a performance profiling tool, we first determined that over 90% of the computation time was spent in the *gradient_refine* subroutine in our global refinement program *emicograd_org*. This subroutine refines the parameters of one particle by minimizing the average phase residual defined in Eq. 1. It does this by varying its orientation and center parameters while the parameters of all other particles remain fixed. This global refinement iterates until the change in phase residual is less than a preset tolerance value. Therefore, the amount of computation in this subroutine varies from particle to particle, depending on the image quality and the accuracy of the initial orientation and center estimates. The next most computationally intensive subroutine computes the phase residual matrix among all particles. This subroutine is first run at the beginning of the refinement to initialize the phase residual matrix and again at the end of program to compute the phase residuals after the refinement. The combined CPU usage of these subroutines accounts for 99% of the total computation for all of our test refinements. Our parallelization algorithm has been designed to reduce the computational cost of these steps.

The computation involved in the global refinement done by *emicograd_org* has data dependencies among all of the particles selected for refinement. This means that in computing the refinement of the parameters of one particle, one must have access to the parameters and residual matrices of all other particles. Consequently, memory usage increases rapidly as the number of particles and resolution increase (Eqs. 2 and 3). There are two possible approaches that provide access to shared data. These are to use either replicated memory or shared memory.

The replicated memory approach duplicates the data to be shared across all processors. This approach was used previously to parallelize a version of global refinement (Johnson et al., 1994; Martino et al., 1994) on a distributed memory machine. However, the data to be shared constitute the bulk of the total memory usage shown in Eqs. 2 and 3, and their replication imposes a limitation on the maximum number of particle images that can be included in the refinement. This limitation could in theory be avoided on distributed memory computers using experimental tools (Keleher et al., 1994), but these are not commercially available.

The shared memory approach does not suffer from the same limitation, because all of the data are stored in a commonly shared memory and can be accessed by processors refining different particles in parallel. Furthermore, programming with shared data is supported by most vendors of shared memory computers. In this analysis we are implicitly assuming that all data are either in RAM or in the system virtual memory. Today, the complexity of programming out-of-core memory management (i.e., using an external disk) makes it prohibitive in terms of code development and maintenance. For problems that are too large to fit in RAM, modern operating system techniques allow users to guide the virtual memory system to achieve performance comparable to out-of-core programming techniques (Park et al., 1996).

Several different types of shared memory architectures are commercially available. We have implemented our global refinement using the shared memory programming model on three of these: an SGI (Silicon Graphics) Power Challenge with a centralized memory system (Fig. 3 *b*), an SGI Origin2000 with distributed shared memory (Fig. 3 *c*), and a KSR-1 (Kendall Square Research) with a cache-only memory system (Kumar et al., 1994) (Fig. 3 *c*). These systems were chosen because they all provide a shared memory programming environment but represent substantially different physical architectures. Their performance comparisons provide useful assessments and guidance for choosing appropriate platforms for future computation. Detailed descriptions of the system architectures and the software packages for these machines are presented in an Appendix, available via the Internet (URL on the World Wide Web: http://www.hpc.uh.edu/~hong/parallel.html or http://ncmi.bcm.tmc.edu). Performance comparisons between the SGI Power Challenge and KSR-1, as well as between static and dynamic scheduling, have been reported elsewhere (Haskell et al., 1997).

The implementations of the parallel code in the shared memory paradigm on the SGI Power Challenge and Origin2000 are identical, and the implementation is similar on the KSR-1. As an illustration, here we only show a segment of our implementation on the SGI Origin2000 machine for the dynamic strategy (Fig. 4). What is shown is the loop indexed by the variable *I_image*. This results in *N_images* separate tasks to be done in parallel. The parallel programming statements are specified in specially formatted comment lines beginning with "c$". Because "c$" is treated as comment lines by sequential Fortran compilers, this parallel code can be directly compiled to run as a sequential program on single processor workstations.

Because the same program segment is executed on each processor, certain variables must have values that are kept private on each processor for the calculation to be performed correctly. These variables are designated as *local* in

```
c$doacross share(L_MOVE,N_SUB_CYCLES,l_l_move,Image)
c$&        share(l_l_move,J_sub_cycle)
c$&        local(I_image,grad,ALine,BLine,
c$&        alpha1,alpha2,Fmod,Fmod1)
c$&        local(RScale, N_summed)
c$MP_SCHEDTYPE=dynamic
c$CHUNK=1
     do I_image = 1, N_images
                refine and update ....
     enddo
```

FIGURE 4 A Fortran program segment illustrating the parallel implementation on the SGI Origin2000 and Power Challenge. Parallel directives typically begin with "c$" and their continuation lines with "c$&". The keywords share and local identify the variables to be shared by multiprocessors and those to be kept local to each processor in the parallel execution of the code. Here a dynamic scheduling type and a chunk size of one iteration are used.

the code to ensure this. The granularity of the computation is defined by the keyword chunk, which is set to be one iteration in the codes, as shown in Fig. 4, and thus the parallelization corresponds to a particle-wise decomposition.

### Parallel performance of global refinement

To evaluate the effectiveness of parallelizing the program, we examine the speedup and efficiency as a function of the number of particle images and processors. Fig. 5 summarizes the parallel global refinement results on the SGI Origin2000 with dynamic scheduling strategies for refining our test sets of 48, 100, and 200 particles. For these data sets, the speedup increases as the number of processors is increased, even up to 24 processors. Despite the fact that the speedup cannot be greater than the number of particles regardless of the number of processors, the efficiency is quite high for all of these tests. Moreover it indicates scalability, because the efficiency increases as the number of particles increases.

Table 1 summarizes the performance and memory requirements on 24 processors of the SGI Power Challenge when various numbers of HSV1 B-capsid particle images are refined. Note that the elapsed time for 355 particles is over 2.5 h. Based on our speedup analysis in Fig. 5, we project that it would take ~2 days on a single processor. It would be impractical to do such a calculation with thousands of particles on a single processor when a higher resolution reconstruction is targeted.

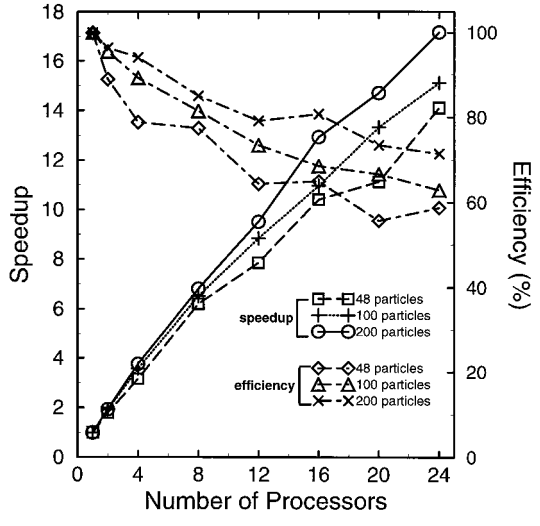


FIGURE 5 Performance of SGI Origin2000. The speedups and parallel efficiencies obtained from refining the 48, 100, and 200 particle test sets with a dynamic scheduling strategy are plotted as a function of the number of processors.

## APPLICATION TO HSV1 B-CAPSID STRUCTURE DETERMINATION

In this section we present our image data and demonstrate the validity of our reconstruction, refined to 13 Å by the parallel refinement procedure. Then we will mention some key structural features appearing at this resolution.

### Data acquisition and processing

Twenty focal pairs of 400-kV micrographs were selected and digitized at 4.67 Å/pixel in a PDS 1010M scanner. Image quality was quantitatively assessed by evaluating CTF rings in the incoherently averaged Fourier transforms of boxed-out particle images, using the CTF determination option of the ICE software system (Zhou et al., 1996). Eight focal pairs, in which the CTF first zero in the close-to-focus

**TABLE 1 Parallel global refinement of different numbers of HSV1 B-capsid particles***

| No. of capsid particles | Memory usage (MB) | Elapsed time (min) | Effective resolution (Å) |
|---|---|---|---|
| 45 | 15.9 | 4.37 | 130 |
| 89 | 35.2 | 13.18 | 50 |
| 178 | 91.9 | 43.70 | 22 |
| 355 | 280.3 | 150.82 | 14[#] |

*A full set of B-capsid images comprising 355 close-to-focus particle images was selected and refined on 24 processors of the SGI Power Challenge. The preliminary orientation and center parameters were determined from the out-of-focus images. This set was then divided into two half-sets (178 particles each), four quarter-sets (89 particles each), and eight small sets (45 particles each), which were independently subjected to global refinement. Resolution was evaluated by computing the phase difference as a function of spatial frequency between independent reconstructions within each set.

[#]The final resolution for the full set of 355 particle images is 13 Å, based on the Fourier ring correlation criterion (see Fig. 7).

micrograph fell within the range 1/10–1/12 $Å^{-1}$ (corresponding to underfocus values of 0.65–1.03 $\mu$m), were chosen for further data processing (Fig. 6). We started with 562 particles boxed out from the selected out-of-focus micrographs (Fig. 1) and were able to determine initial orientation and center parameters for 95% of them. These initial orientation parameters were used as the starting values for processing the corresponding particles from the close-to-focus micrographs. During data collection with the spot-scan procedure, the hysteresis of the electron beam makes it difficult to scan exactly the same specimen locations in a focal pair. Thus some of the 532 particle images in the close-to-focus micrographs were incomplete and were discarded, resulting in only 355 particles being used for extensive local and global refinement, and for subsequent 3D reconstruction.

## Evaluation of reconstructions

The full data set of 355 B-capsid particles was refined by using a combination of local and parallel global refinement (Fig. 1) to a nominal resolution of 12.3 Å. At this resolution, the averaged cross common-line phase residual between raw particle images and computed projections is 51°. This full data set was then arbitrarily divided into two subsets from which independent reconstructions were calculated, using data truncated at ~12.3 Å. The effective resolution of the final map, based on a phase difference of 45° between the two independent reconstructions, is ~14–15 Å (Fig. 7). This choice of phase difference as a resolution cutoff is reasonable, because a random phase difference would be 90°. If we use the less stringent Fourier ring correlation criterion, the effective resolution of the final reconstruction is ~13 Å (Fig. 7).

The practical number of particle images needed for a 3D reconstruction at a given resolution depends on many experimental variables and thus is generally unknown. One way to obtain an indication of the sufficiency of the particle numbers is to examine the resolution of the reconstruction as a function of the number of particles used. Table 1 summarizes the resolutions attained when various numbers of HSV1 B-capsid particle images are refined. Increasing the number of particles greatly improves the resolution (Fig. 7), and the *S*/*N* ratio of the computed 3D reconstructions (Fig. 8 *a–d*).

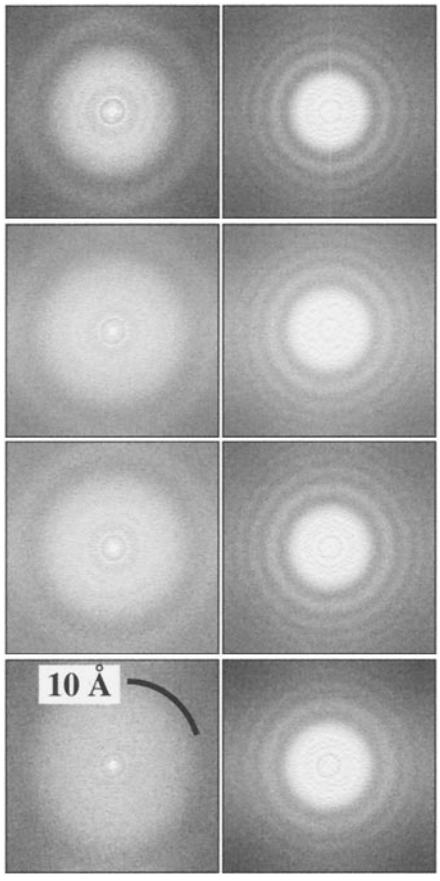


FIGURE 6 Incoherent averages of Fourier transforms computed from 400-kV spot-scan images of ice-embedded HSV1 B-capsids in four representative focal pairs of close-to-focus (*left*) and out-of-focus (*right*) images. The rings represent the contrast transfer functions from which the defocus values can be determined and the maximum potentially attainable resolution in each of the data sets. Only the close-to-focus data (*left*) are used in the final 3D reconstructions. The arc in the bottom left spectrum marks the CTF first zero, which occurs at spatial frequency 1/10 $Å^{-1}$.

## HSV1 B-capsid structure at 13 Å

The 13-Å structure reconstructed from the full data set is shown in Fig. 9. The maps are displayed with an isosurface value of 1$\sigma$ (standard deviation) from the average density, which is near zero. To facilitate the interpretation of the

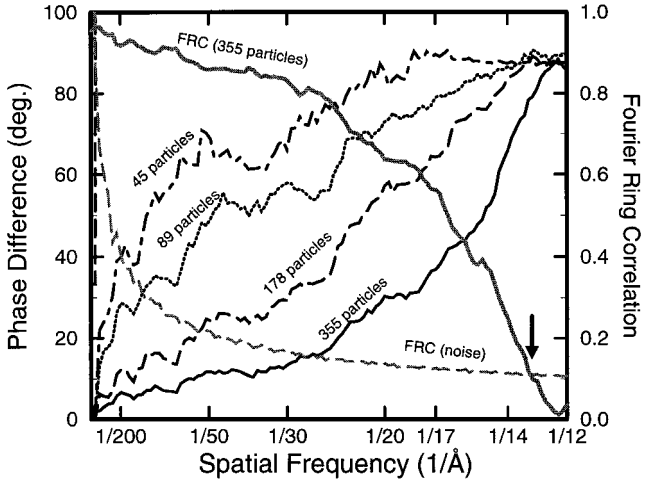


FIGURE 7 Resolution evaluation of the reconstructions obtained by global refinement of different numbers of particles. The phase differences between independent reconstructions from equally sized subsets are plotted as a function of spatial frequency in solid (355 particles), dashed (178 particles), dotted (89 particles), and dotted-dash (45 particles) curves. The Fourier ring correlation (FRC) between two independent reconstructions of the 355-particle set is also shown (*solid gray*); this intersects at ~13 Å with the curve plotting twice the expected FRC for Gaussian noise (*dashed gray*).

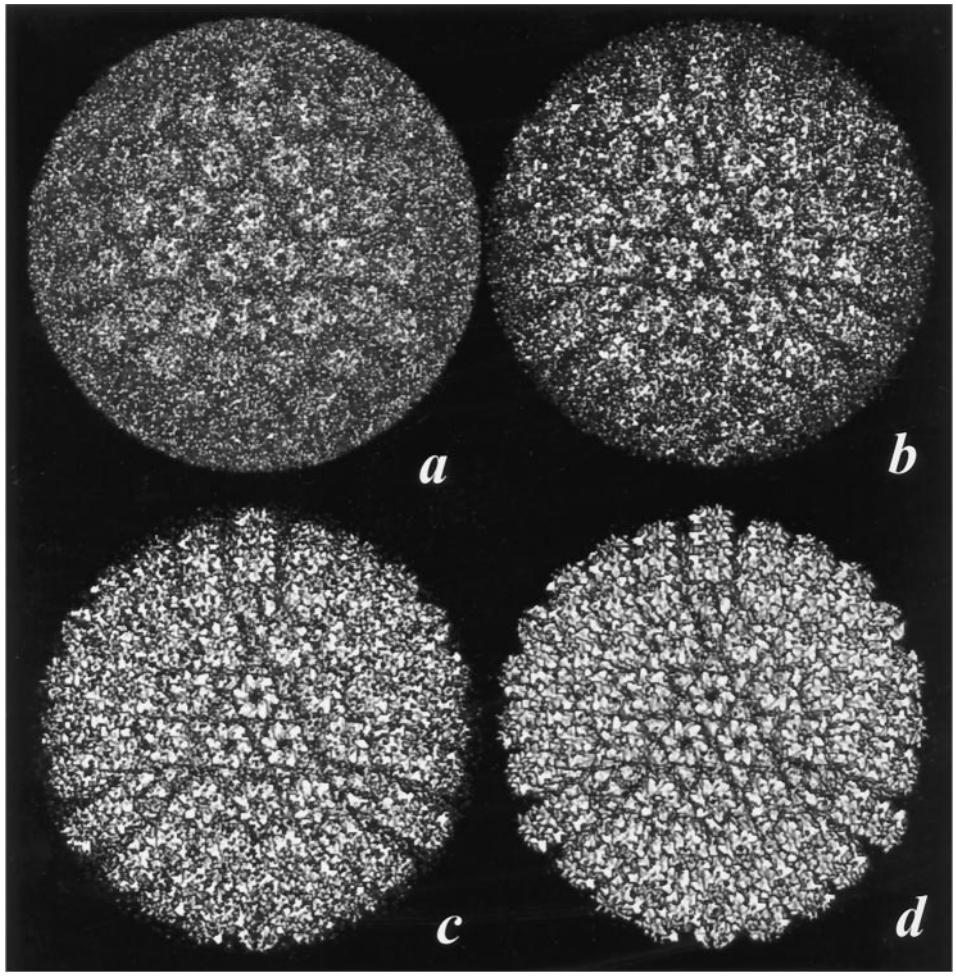


FIGURE 8 Surface representations of the 3D structure of HSV1 B-capsid generated from 45 particles (*a*), 89 particles (*b*), 179 particles (*c*), and 355 particles (*d*), independently refined by using global refinement on the SGI Power Challenge with 24 processors.

complex $T = 16$ structure, a segment of capsid including all the components of an asymmetrical unit is highlighted in Fig. 9 *a*, and the three types of morphological unit (penton, hexon, and triplex) are colored to facilitate their identification. Fig. 9 *b* shows a schematic diagram of one of the 20 icosahedral triangular facets of the $T = 16$ capsid, which contains three asymmetrical units. Each asymmetrical unit consists of one-fifth of the penton, one P-hexon, one C-hexon, one-half E-hexon, one each of the five $T_a$-$T_e$ triplexes, and one-third of the $T_f$ triplex. The 16 quasi-equivalent capsomere subunits and six types of triplex in a single asymmetrical unit are highlighted according to the same color codes as in Fig. 9 *a*. Each of the five subunits of the penton consists of one molecule of the 150-kDa protein VP5 (Steven and Spear, 1997; Zhou et al., 1994). By difference imaging of wild-type and VP26-minus recombinant capsids, each hexon has been shown to consist of six copies each of VP5 and VP26 (Trus et al., 1995; Zhou et al., 1995). Each of the triplexes is a heterotrimer made of one copy of VP19c and two copies of VP23 (Newcomb et al., 1993).

To reveal the molecular interactions in detail, a contiguous group of morphological units has been extracted computationally, using modules custom-designed for the graphics software Explorer (NAG, Inc.). It is viewed from the outside (Fig. 9 *c*) and inside (Fig. 9 *d*) of the capsid. At 13-Å resolution, the mass density distributions of the VP5 subunits and the triplexes are better resolved than those shown previously in the 25-Å map (Zhou et al., 1994). The separations of the subunits in both penton and hexons can be discerned, particularly on their upper and middle domains (Fig. 9 *c*), but this can be difficult, for example, within the capsid floor, where the bundles of density intertwine (Fig. 9 *d*). The triplexes maintain an extensive network of lateral interactions with adjacent penton/hexon subunits. The pattern of interactions varies, as indicated by arrowheads in Fig. 9 *c* and schematized in Fig. 9 *b*. Further structural details of the protein subunits can be seen at our web sites.

## DISCUSSION

### Electron cryomicroscopy and computer reconstruction of icosahedral particles

The electron imaging technique for ice-embedded icosahedral particles has been improved significantly in the past few years by a combination of factors, including improvements in the mechanical and thermal stability of cryoholders, reductions in beam-induced specimen movement using

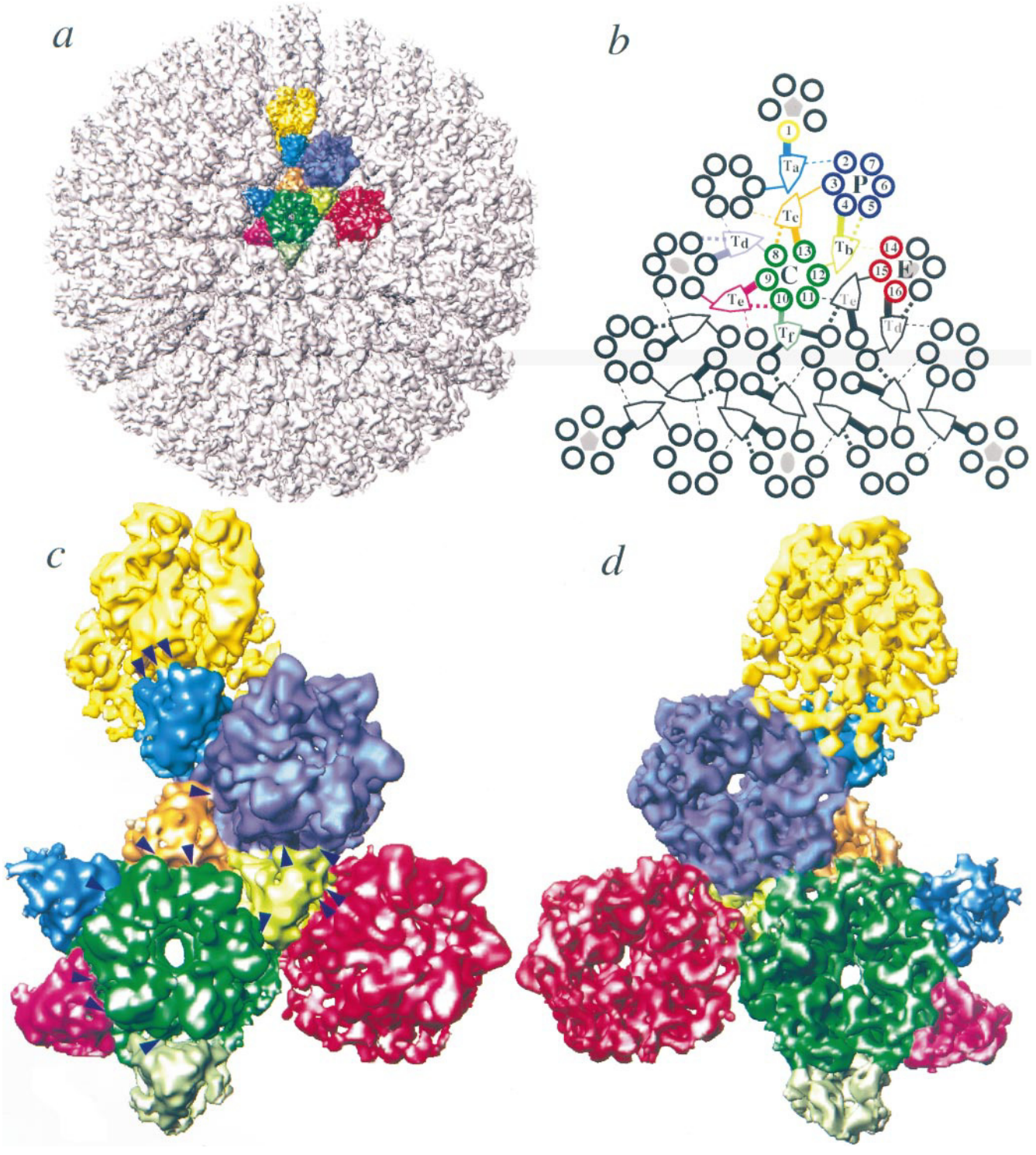


FIGURE 9 3D structure of the HSV1 B-capsid refined to 13-Å resolution by parallel global refinement. (*a*) The whole capsid is shown as a shaded surface with the structural components in one asymmetrical unit in color. Included in an asymmetrical unit are morphological components from penton (*yellow*), P-hexon (*blue*), E-hexon (*red*), C-hexon (*green*), and six different types of triplex (*other colors*). (*b*) The schematic diagram of one triangular face of the herpesvirus capsid illustrates the interactions between the subunits (VP5) of penton and hexons (1–16) and triplexes $T_a$, $T_b$, $T_c$, $T_d$, $T_e$, and $T_f$. Subunits within an asymmetrical unit are shown with the same color coding as in *a*. Four kinds of interactions between triplex and VP5 are depicted, including strong head (*thick solid line*), weak head (*thick dashed line*), tail (*thin solid line*), and arm (*thin dashed line*). (*c*) The contiguous group of morphological components colored in *a* is blown up and shown as viewed from outside (*c*) and inside (*d*). Arrowheads in *c* indicate the contacts between triplexes and their adjacent penton and hexons. Only those visible at this view are indicated.

spot-scan procedures, and improved coherence of the electron source. We used the 400-kV electron cryomicroscope for data collection because of its high-resolution performance (Brink and Chiu, 1991) and its relatively large depth of field (Zhou and Chiu, 1993), which is required for achieving the highest possible resolutions for a large particle such as the herpesvirus capsid. A disadvantage of this cryomicroscope is the lack of a field emission gun that would improve the spatial coherence of the electron beam. Thus our experimental approach is to record images close to focus, so that the first zero of the CTF extends to as high a resolution as possible, while still allowing the particle to be identified, centered, and analyzed for 3D reconstruction. Indeed, the images used in our analysis have the CTF first zeros at frequencies higher than those reported in any previous work on icosahedral particles. This was done to optimize the recording of higher resolution data by minimizing the damping of the high-resolution information caused by the limited spatial coherence of the incident electron beam. Such a constraint can be relaxed if a field emission gun with high spatial and temporal coherence is employed (Böttcher et al., 1997; Conway et al., 1997; Trus et al., 1997). However, the techniques developed here would still apply to give improved resolution.

Fig. 6 shows the incoherent averages of computed particle diffraction patterns from four of eight focal pairs of 400-kV spot-scan images recorded using a conventional $LaB_6$ electron source. The quality of these data is evident from the isotropic contrast seen beyond the first zero of the CTF at ~1/12 to 1/10 $Å^{-1}$. In this study we have not applied any correction to retrieve data beyond the CTF first zero.

Our data-processing scheme is the result of combining and modifying steps used by various investigators (Böttcher and Crowther, 1996; Fuller, 1987; Zhou et al., 1995). The major features of our scheme as outlined in Fig. 1 are as follows. First, we rely on the out-of-focus particle images to determine the center and orientation estimates for the corresponding close-to-focus particle images that are used for the eventual reconstruction (Cheng et al., 1992). Next, a local refinement, based on projections of preliminary 3D maps, is used iteratively with a global refinement that does not rely on model projections. Our refinement is based on the principle of minimizing the average phase residuals along the self and cross common-lines among many particle images (Böttcher and Crowther, 1996; Fuller, 1987; Zhou et al., 1995).

Before undertaking this analysis, we did not know whether image collection and data-processing techniques developed to date could adequately handle images of increased complexity from which higher resolution results are desired. Parallelizing a program requires extensive modifications that may uncover flaws that only appear when the program capabilities are pushed to their limits, e.g., numerical accuracy and data set size. The fact that the current map displays the prominent features of the HSV1 capsid previously seen in the low-resolution map reinforces our confidence in the correctness of the modified code.

An alternative method for determining the orientation parameters of icosahedral particles has been developed that involves searching for a cross-correlation match between the raw particle images and a series of evenly sampled 3D model projections (Baker and Cheng, 1996). However, there has been no rigorous comparison of the two methods using the same data sets to determine their relative accuracy and efficiency.

### Reconstruction using parallel processing

Parallelization has been used previously to process icosahedral virus particle images (Johnson et al., 1994; Martino et al., 1994). In this study, the initial orientation search was also parallelized, and the implementation was carried out on the Intel iPSC/860, a distributed memory computer (Fig. 3 *a*). The initial orientation search is an early step in our analysis that generates a list of possible orientations for each particle by minimizing the self common-line phase residuals (Fig. 1). The amount of computation required for the initial orientation search for a single particle is now considered to be too small to require parallel processing (~10 s CPU time on the SGI R10000 and IBM RS6000). For the large numbers of particles used in current studies, the initial orientation search for each particle is best treated as a separate problem. Because there are no data dependencies among different particles in the initial orientation search program, parallelism can be achieved automatically by executing multiple searches in parallel. For instance, our initial orientation search program *ortAll* creates separate processes, one for each particle, for different processors to execute independently (Zhou, 1995). This could be done in a similar way on a distributed memory machine such as the iPSC/860.

A limitation of the approach used by Johnson et al. (Johnson et al., 1994; Martino et al., 1994) for parallelizing the global refinement step is the use of replicated memory, the drawbacks of which were discussed earlier (Parallel Implementation of Global Refinement). We have avoided this limitation by using shared memory.

## CONCLUSIONS

Electron cryomicroscopy is a technique whose role in structural virology continues to grow. Although high-quality experimental image data can be readily obtained, to make the approach more powerful and accessible, it is necessary to improve the computational tools used to retrieve the 3D structure of large-size particles such as those over 1000 Å in diameter. We have adopted an image-processing scheme that uses focal-pair image data and involves iterative steps of local and global refinement to determine the orientations and centers of the particle images.

The global refinement is the most computationally intensive step in this process and is well suited for execution on

computers with multiple processors operated in shared-memory mode. We have implemented this refinement on SGI and KSR parallel computers. The parallel refinement program makes it routine to refine the hundreds of very noisy particle images that must be averaged to obtain sufficient statistical definition of higher resolution 3D structural features. We have applied this method to the structural determination of the herpesvirus capsid to 13-Å resolution. This map reveals new structural features of the subunits and their interactions.

We thank Mr. Matt Dougherty for help with graphics and Dr. Pamela Thuman-Commike for helpful discussions about the manuscript. We appreciate many helpful comments made by the reviewers.

This project has been supported by grants from the National Institutes of Health (RR02250, AI38469), the National Library of Medicine (LM07093), the National Science Foundation (BIR-9412521, BIR-9413229), the Human Frontier Science Program (RG-537/96), and the W. M. Keck Foundation.

## REFERENCES

Baker, T. S., and R. H. Cheng. 1996. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J. Struct. Biol.* 116:120–130.

Baker, T. S., J. Drak, and M. Bina. 1988. Reconstruction of the three-dimensional structure of simian virus 40 and visualization of chromatin core. *Proc. Natl. Acad. Sci. USA.* 85:422–426.

Baker, T. S., and J. E. Johnson. 1997. Principles of virus structure determination. *In* Structural Biology of Viruses. Oxford University Press, New York. 38–79.

Böttcher, B., and R. A. Crowther. 1996. Difference imaging reveals ordered regions of RNA in turnip yellow mosaic virus. *Structure.* 4:387–394.

Böttcher, B., S. A. Wynne, and R. A. Crowther. 1997. Detemination of the fold of hepatitis B virus core protein by electron cryo-microscopy. *Nature.* 386:88–91.

Brent, R. P. 1973. Algorithms for Minimization without Derivatives. Prentice-Hall, Englewood Cliffs, NJ.

Brink, J., and W. Chiu. 1991. Contrast analysis of cryo-images of *n*-paraffin recorded at 400 kV out to 2.1 Å resolution. *J. Microsc.* 161: 279–295.

Casjens, S. 1997. Principles of virion structure, function and assembly. *In* Structural Biology of Viruses. Oxford University Press, New York. 3–37.

Caspar, D. L. D., and A. Klug. 1962. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27:1–32.

Cheng, R. H., N. H. Olson, and T. S. Baker. 1992. Cauliflower mosaic virus: a 420 subunit ($T = 7$), multilayer structure. *Virology.* 186: 655–668.

Cheng, R. H., V. S. Reddy, N. H. Olson, A. J. Fisher, T. S. Baker, and J. E. Johnson. 1994. Functional implications of quasi-equivalence in a $T = 3$ icosahedral animal virus established by cryo-electron microscopy and x-ray crystallography. *Structure.* 2:271–282.

Chiu, W., T.-W. Jeng, L. L. Degn, and B. V. Prasad. 1986. Potential for high-resolution electron crystallography at intermediate high voltage. *Ann. N.Y. Acad. Sci.* 483:149–156.

Chiu, W., and T. J. Smith. 1994. Structural studies of virus-antibody complexes by electron cryomicroscopy and x-ray crystallography. *Curr. Opin. Struct. Biol.* 4:219–224.

Conway, J. F., N. Cheng, A. Zlotnick, P. T. Wingfield, S. J. Stahl, and A. C. Steven. 1997. Visualization of the 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature.* 386:91–94.

Conway, J. F., R. L. Duda, R. W. Hendrix, and A. C. Steven. 1995. Proteolytic and conformational control of virus capsid maturation: the bacteriophage HK97 system. *J. Mol. Biol.* 253:86–99.

Crowther, R. A. 1971. Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. *Philos. Trans. R. Soc. Lond. B.* 261:221–230.

Crowther, R. A., D. J. DeRosier, and A. Klug. 1970. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Proc. R. Soc. Lond.* 317:319–340.

DeRosier, D. J., and A. Klug. 1968. Reconstruction of three-dimensional structures from electron micrographs. *Nature.* 217:130–134.

Downing, K. H. 1991. Spot-scan imaging in transmission electron microscopy. *Science.* 251:53–59.

Dryden, K. A., G. J. Wang, M. Yeager, M. L. Nibert, K. M. Coombs, D. B. Furlong, B. N. Fields, and T. S. Baker. 1993. Early steps in reovirus infection are associated with dramatic changes in supramolecular structure and protein conformations. *J. Cell Biol.* 122:1023–1041.

Fuller, S. D. 1987. The $T = 4$ envelope of Sindbis virus is organized by interactions with a complementary $T = 3$ capsid. *Cell.* 48:923–934.

Fuller, S. D., S. J. Butcher, R. H. Cheng, and T. S. Baker. 1996. Three-dimensional reconstruction of icosahedral particles—the uncommon line. *J. Struct. Biol.* 116:48–55.

Harrison, S., D. C. Wiley, and J. J. Skehel. 1996. Virus structure. *In* Virology. Lippincott-Raven, Philadelphia. 59–100.

Haskell, K. H., Z. H. Zhou, J. H Spears, W. Chiu, and L. R. Scott. 1997. Improved 3D reconstruction of virus structure through parallel processing. *In* Eighth SIAM Conference on Parallel Processing. SIAM, Minneapolis, MN.

Johnson, C. A., N. I. Weisenfeld, J. F. C. B. L. Trus, R. L. Martino, and A. C. Steven. 1994. Orientation determination in the 3D reconstruction of icosahedral viruses using a parallel computer. *In* Supercomputing '94. IEEE Computer Society Press, Los Alamitos, CA. 550–559.

Johnson, J. E. 1996. Functional implications of protein-protein interactions in icosahedral viruses. *Proc. Natl. Acad. Sci. USA.* 93:27–33.

Keleher, P., S. Dwarkadas, A. L. Cox, and W. Zwaenepoel. 1994. TreadMarks: distributed shared memory on standard workstations and operating systems. *In* Winter 94 usenix conference. 115–131.

Kumar, V., A. Grama, A. Gupta, and G. Karypis. 1994. Introduction to Parallel Computing. Benjamin/Cummings Publishing Co., Redwood City, CA.

Lewis, T. G., and H. El-Rewini. 1992. Introduction to Parallel Computing. Prentice-Hall, Englewood Cliffs, NJ.

Martino, R. L., C. A. Johnson, E. B. Suh, B. L. Trus, and T. K. Yap. 1994. Parallel computing in biomedical research. *Science.* 265:902–908.

Marvik, O. J., T. Dokland, R. H. Nokling, E. Jacobsen, T. Larsen, and B. H. Lindqvist. 1995. The capsid size-determining protein Sid forms an external saffold on phage P4 procapsids. *J. Mol. Biol.* 251:59–75.

McKenna, R., D. Xia, P. Willingmann, L. L. Ilag, S. Krishnaswamy, M. G. Rossmann, N. H. Olson, T. S. Baker, and N. L. Incardona. 1992. Atomic structure of single-stranded DNA bacteriophage $\phi$X174 and its functional implications. *Nature.* 355:137–143.

Newcomb, W. W., B. L. Trus, F. P. Booy, A. C. Steven, J. S. Wall, and J. C. Brown. 1993. Structure of the herpes simplex virus capsid: molecular composition of the pentons and the triplexes. *J. Mol. Biol.* 232: 499–511.

Olson, N. H., and T. S. Baker. 1989. Magnification calibration and the determination of spherical virus diameters using cryo-microscopy. *Ultramicroscopy.* 30:281–97.

Olson, N. H., P. R. Kolatkar, M. A. Oliveira, R. H. Cheng, J. M. Greve, A. McClelland, T. S. Baker, and M. G. Rossmann. 1993. Structure of a human rhinovirus complexed with its receptor molecule. *Proc. Natl. Acad. Sci. USA.* 90:507–511.

Park, Y., R. Scott, and S. Sechrest. 1996. Virtal memory versus file interfaces for large, memory-intensive scientific applications. *In* Supercomputing '96. IEEE Computer Society Press, Los Alamitos, CA.

Prasad, B. V. V., P. E. Prevelige, E. Marietta, R. O. Chen, D. Thomas, J. King, and W. Chiu. 1993. Three-dimensional transformation of capsids associated with genome packaging in a bacterial virus. *J. Mol. Biol.* 231:65–74.

Prasad, B. V. V., R. Rothnagel, C. Q.-Y. Zeng, J. Jakana, J. A. Lawton, W. Chiu, and M. K. Estes. 1996. Visualization of ordered genomic RNA and localization of the transcription enzymes in rotavirus. *Nature.* 382: 471–473.

Prasad, B. V. V., G. J. Wang, J. P. M. Clerx, and W. Chiu. 1988. Three-dimensional structure of rotavirus. *J. Mol. Biol.* 199:269–275.

Smith, T. J., A. G. Mosser, and T. S. Baker. 1995. Structural studies on the mechanisms of antibody-mediated neutralization of human rhninovirus. *Semin. Virol.* 6:233–242.

Steven, A. C., and P. G. Spear. 1997. Herpesvirus capsid assembly and enelopment. *In* Structural Biology of Viruses. Oxford University Press, New York. 312–351.

Stewart, P. L., S. D. Fuller, and R. M. Burnett. 1993. Difference imaging of adenovirus: bridging the resolution gap between X-ray crystallography and electron microscopy. *EMBO J.* 12:2589–2599.

Thuman-Commike, B. G. P. A., J. A. Malinski, J. King, and W. Chiu. 1997. Role of the scaffolding protein in P22 procapsid size determination suggested by $T = 4$ and $T = 7$ procapsid structures. *Biophys. J.* 73:559–568.

Thuman-Commike, P. A., and W. Chiu. 1996. PTOOL: a software package for the selection of particles from electron cryomicroscopy spot-scan images. *J. Struct. Biol.* 116:41–47.

Thuman-Commike, P. A., and W. Chiu. 1997. Improved common-line based icosahedral virus particle image orientation estimation algorithms. *Ultramicroscopy.* 68:231–256.

Trus, B. L., F. L. Homa, F. P. Booy, W. W. Newcomb, D. R. Thomsen, N. Cheng, J. C. Brown, and A. C. Steven. 1995. Herpes simplex virus capsids assembled in insect cells infected with recombinant baculoviruses: structural authenticity and localization of vp26. *J. Virol.* 69:7362–7366.

Trus, B. L., R. B. Roden, H. L. Greenstone, M. Vrhel, J. Schiller, and F. P. Booy. 1997. Novel structural features of bovine papillomavirus capsid revealed by a three dimensional reconstruction to 9 Å resolution. *Nature Struct. Biol.* 4:411–418.

Wilson, G. V. 1995. Practical Parallel Programming. MIT Press, Cambridge, MA.

Zhou, Z. H. 1995. High-Resolution Three-Dimensional Electron Cryomicroscopy and Reconstruction of Herpes Simplex Virus Capsids. Baylor College of Medicine, Houston, TX.

Zhou, Z. H., and W. Chiu. 1993. Prospects for using an IVEM with a FEG for imaging macromolecules towards atomic resolution. *Ultramicroscopy.* 49:407–416.

Zhou, Z. H., S. Hardt, B. Wang, M. B. Sherman, J. Jakana, and W. Chiu. 1996. CTF determination of images of ice-embedded single particles using a graphics interface. *J. Struct. Biol.* 116:216–222.

Zhou, Z. H., J. He, J. Jakana, J. Tatman, F. Rixon, and W. Chiu. 1995. Assembly of VP26 in HSV-1 inferred from structures of wild-type and recombinant capsids. *Nature Struct. Biol.* 2:1026–1030.

Zhou, Z. H., B. V. V. Prasad, J. Jakana, F. Rixon, and W. Chiu. 1994. Protein subunit structures in the herpes simplex virus A-capsid determined from 400 kV spot-scan electron cryomicroscopy. *J. Mol. Biol.* 242:458–469.