The Suitability of a Monofunctional Reagent of an Undecagold Cluster for Phasing Data Collected from the Large Ribosomal Subunits from *Bacillus stearothermophilus*

An electron density map of the large ribosomal subunit from *Bacillus stearothermophilus* was obtained at 2.6 Å resolution by single isomorphous replacement (SIR) from a derivative formed by specific quantitative labeling with a dense undecagold cluster. For derivatization, a monofunctional reagent of this cluster was bound to a sulfhydryl group of a purified ribosomal protein, which was in turn reconstituted with core particles of a mutant lacking this protein. The native, mutated, and derivatized 50S ribosomal subunits crystallize under the same conditions in the
same space group. Under favorable conditions, crystals of the derivatized subunit proved to be isomorphous with the native ones, whereas the crystals of the mutant may have somewhat different packing.

After resolving the SIR phase ambiguity by solvent flattening, the electron density shows a packing that is consistent with the noncrystallographic symmetry found by Patterson searches as well as with the motif observed in electron micrographs of thin sections of the crystals. These studies established that phase information can be obtained from heavy metal clusters, even when the crystals under investigation are unstable and weakly diffracting. These results encouraged further effort at the construction of specifically derivatized crystals from other ribosomal particles that diffract to higher resolution. © 1995 John Wiley & Sons, Inc.

INTRODUCTION

Ribosomes provide the site for the translation of the genetic code into polypeptide chains in all living cells. They are built of two independent subunits of unequal size that associate upon the initiation of protein biosynthesis. A typical bacterial ribosome (70S) has a molecular weight of 2.3 million daltons and its large (50S) and small (30S) subunits have molecular weights of 1.45 and 0.85 million daltons, respectively. About one third of its mass is comprised of 58–73 different proteins, depending on the source. The rest is composed of three rRNA chains with a total of about 4500 nucleotides.

Crystals diffracting best to 2.9 Å have been grown from active ribosomal particles from halophilic and thermophilic bacteria.1 X-ray crystallographic data of reasonable quality are being collected with synchrotron radiation at cryo-temperature from ribosomal particles of the wild type and mutated bacteria, chemically modified particles, and complexes of ribosomal particles with components of protein biosynthesis (mRNA, tRNA, and nascent protein chains).

In this article we describe our attempts at low-resolution phasing of the diffraction data obtained from the monoclinic crystals of the large ribosomal subunits from Bacillus stearothermophilus (BSOS). These particles were chosen although their crystals diffract to low resolution, because (a) their biochemistry and genetics are rather well characterized to allow the manipulations required for their specific derivatization; (b) they possess an exposed -SH to which an undecagold cluster could be quantitatively bound2; and (c) because a low resolution model, which may be exploited during the course of structure determination, has been reconstructed from tilt series of electron micrographs taken from their negatively stained two-dimensional arrays.3

PHASING IN MACROMOLECULAR CRYSTALLOGRAPHY

In macromolecular crystallography, the assignment of phases to the observed structure factor amplitude is the most crucial, albeit most complicated step in structure determination. Since the phases cannot be directly measured, their elucidation remains the most complicated and least predictable tasks in structure determination, even for average-size proteins, of molecular weights of 15–80 Kda. Clearly, for ribosomal crystals, the magnitude and the complexity of this step is greatly enhanced.

The commonly used method for the determination of the phases of new structures is isomorphous replacement. It is based on the changes in the structure factor amplitudes that are caused by the addition of heavy atoms to the native compound. It requires attachment of heavy atoms at a limited number of sites within the unit cell, while keeping the crystal structure isomorphous to that of the native molecule. When only one derivative is available, the single isomorphous replacement (SIR) procedure, which often yields ambiguous phasing, is used. When more than one such derivative can be obtained, multiple isomorphous replacement is used. In this procedure, the information from the different derivatives is combined, leading to more reliable phasing.

The derivatization reagents are chosen according to their potential ability to induce measurable signals. For proteins of average size (e.g., with molecular weights of 15–80 Kda, useful heavy atom derivatives consist of one or two heavy-metal atoms. Because of the large size of the ribosome, ideal compounds for derivatization are compact and dense materials of a proportionally larger number of heavy atoms. Therefore multimetal compounds, such as polyheteroanions (e.g., K14NaP5W3001 or dense metal clusters, such as undecagold (Figure 1) are being used by us for soaking and for specific binding, respectively.

Heavy-atom derivatives of crystals of biological
Suitability of a Monofunctional Reagent

The monofunctional reagent of the undecagold cluster used for the derivatization.

Macromolecules are routinely obtained by soaking crystals in solutions of the heavy-atom compounds or by cocrystallization of the macromolecule together with the heavy atom. Using these equilibrium procedures, productive derivatization in one or a few sites is largely a matter of chance, but the probability of obtaining a useful derivative for a typical macromolecule is sufficiently high that more sophisticated techniques are rarely needed.

**SPECIFIC LABELING BY AN UNDECAGOLD CLUSTER**

The undecagold cluster (GC) used for derivatization contains a core of 8.2 Å diameter, consisting of eleven gold atoms linked directly to each other, surrounded by hydrophilic organic groups forming a unique, well-defined compound with an overall diameter of 22 Å. This cluster, of a molecular weight of 6.2 Kda, combines high electron density with water solubility, and could be prepared as a monofunctional reagent, aiming first at binding to exposed free sulfhydryls. Since the gold cluster is rather bulky, accessibility was enhanced by attaching a maleimido or an iodoacetyl group to the cluster through a short aliphatic chain (Figure 1). The latter was designed to avoid the chirality, which may be introduced by the reaction of the sulfhydryl with the double bond of the maleimido moiety.

Assuming reasonable isomorphism and a full single-site occupancy of the undecagold cluster, the average differences in diffraction intensities between native and GC-bound 50S particles is expected to be about 15%. Hence, provided these requirements are fulfilled, the undecagold cluster is likely to induce measurable differences between the intensities of the reflections of the native crystals and those of the derivatized ones. As the core of the undecagold cluster is about 8.2 Å in diameter, it can be treated as a single scattering group at low to intermediate resolution.

A mutant lacking one ribosomal protein, BL11, was obtained by the addition of the antibiotic thiostrepton to the bacterial growth medium. In parallel, protein BL11 was isolated from native particles. The mere fact that the bacteria can grow without protein BL11, alongside our ability to crystallize the cores lacking this protein under the same conditions as used for the native particles, indicate that the removal of this protein does not cause gross conformational changes in the ribosome, and that protein BL11 it is not likely to be involved in crystal packing. The monofunctional reagent of the undecagold cluster was bound to protein BL11, by reacting with its only cysteine. Al-
though the binding of the cluster was carried out under denaturing conditions, and although the molecular weight of the undecagold cluster (6.2 Kda) approaches half of that of the protein (15.5 Kda), the derivatized protein could be reconstituted into cores of mutated ribosomes lacking BL11, yielding fully derivatized particles.

**SIMULATION STUDIES ON THE PHASING POWER OF THE UNDECAGOLD CLUSTER**

In order to assess the phasing power of the gold cluster and the influence of various parameters on the phasing process, model calculations have been performed. The envelope of the reconstructed image mentioned above was filled uniformly with a random distribution of about 115,000 nonhydrogen atoms, according to the known chemical composition of the large ribosomal subunit of *Escherichia coli*.

Assuming the existence of more than one particle in the asymmetric unit, and still not knowing the correct number (which was found later), two such models were placed in the asymmetric unit, keeping the dimensions and space group of B50S. The solvent was modeled as a uniform density outside the molecule convoluted with a Gaussian using a temperature factor of 320 Å². The solvent density was set to its experimentally determined value of 0.36 electrons per cubic Å. Structure factors were calculated and an error model mimicking the experimental R-factor distributions was applied to yield a native data set. A data set with simulated lack of isomorphism was generated by changing the p angle of the unit cell by 0.5°. Derivative data sets were constructed by placing the gold cluster into the native structure in a nonoverlapping fashion.

To monitor the phasing power we chose the variable $|F_{\text{native}}^2 - F_{\text{derivative}}^2|/F_{\text{native}}^2$, where $F_{\text{native}}$ is the native and $F_{\text{derivative}}$ the derivative structure factor. Figure 2 shows the dependence of this variable on the Bragg resolution. A significant signal appears at around 50 Å, increases steadily until 8 Å, which is about the diameter of the gold cluster core, and decreases at higher resolution. This reflects the fact that a group of atoms can only be considered as a single group scatterer up to a resolution close to its diameter. At resolutions higher than that, individual atoms rather than group properties determine the scattering behavior.

Among the parameters checked, the occupancy and the mean displacement of the gold-cluster position were found to be the most influential. The influence of the occupancy is practically linear. The influence of the mean displacement can be seen in Figure 2. As seen, mean displacement of 5 Å (roughly half the cluster diameter) generated a significant loss in signal.

The analysis of the simulated difference Patterson maps emphasized the importance of two additional parameters: experimental error and lack of isomorphism. Figure 3a shows a clear peak at the positions of the gold cluster in the Harker section at 2x, 0, 2z generated from error-free data assuming a mean displacement of the cluster of 3 Å. Convoluting the data with a Gaussian error of 5% leads to a Patterson map in which the signal is threefold lower, but still quite clear. Figure 3b shows the same section, calculated from simulated data assuming lack of isomorphism of 0.5° difference in p. It is evident that the noise level is rather high, and the correct peaks have a signal of only 2.4% of the origin peak. Still, the correct positions could be retrieved using vector verification.

These studies indicated that it should be possible to obtain useful phase information between 50 and 8 Å from the undecagold cluster, provided the mean displacement of the cluster position, the experimental error and the lack of isomorphism could be kept relatively low. High occupancy has been ensured by the way the derivative was constructed.

**THE CRYSTALLOGRAPHIC DATA**

The crystals of native, mutated, and derivatized B50S were grown according to Ref. 8 with small
modifications. They appear as monoclinic needles with maximum dimensions of 0.8 × 0.2 × 0.04 mm³. Their space group is C2, with unit cell dimensions of 300 × 547 × 377 (±1%) Å and β = 111.9°. Their extremely weak diffraction power and high sensitivity to the X-radiation dictate crystallographic data collection with intense synchrotron radiation at cryo-temperature.

Data sets of four native, four derivatized, and one mutated crystals have been evaluated (Table I). Two of these sets (41 and 53) were collected from the same crystal in two measuring periods, separated by an interval of five months, during which the crystal was stored in a solid propane-liquid nitrogen dewar at 85 K. The maximum usable data available from some of the crystals extends to 15 Å, although occasionally the diffraction extended to higher resolution (11 Å). The crystals suffer from significant nonisotropic mosaicity, 0.3°-1.0°, depending also on the resolution. In several cases, the variations in cell constants were over 1%, indicating lack of isomorphism, which was verified in unsuccessful attempts to merge data from more than one crystal. Noteworthy is the deviation of the volume of the mutant from those of the native and derivatized ones.

**DIRECT OBSERVATION OF CRYSTAL PACKING**

The large size of the ribosomal particles enables their observation by electron microscopy. Three-dimensional crystals are embedded in epon, sectioned at preferred orientations in extremely thin slices, of a width similar to that of a single unit cell, and positively stained with uranyl-acetate. Visual inspection and symmetry considerations, were then used for the determination of the packing motifs of the particles in the unit cells and revealed internal arrangement in triplets (Figure 4). This provided a valuable hint, since previous attempts at measuring or calculating the density of the B50S crystals, indicated only that the asymmetric unit contains 2–4 particles, but could not lead to conclusive determination of the actual number.

**SELF-ROTATION SEARCHES**

The existence of three particles in the asymmetric unit was demonstrated conclusively by a self-rotation search that was used for locating the noncrystallographic symmetry elements in the crystallographic asymmetric unit. These searches were performed using data collected from native (46), mutant lacking protein L11 (83), and gold cluster derivative (41) (Table I). At resolution range 50–20 Å, with a Patterson cutoff of 150 Å, a clear peak (52% of the origin) was found in the κ = 120° section, denoting a threefold noncrystallographic axis at the spherical polar angles (168, 90, 120). This threefold axis persists in other resolution ranges (i.e., 50–30, 70–30, and 40–20 Å) and with different Patterson cutoff radii (140–180 Å). Besides, a clear (50% of the origin) twofold noncrystallographic peak was found in the κ = 180° section of

**FIGURE 3** The influence of the lack of isomorphism (b) on the signal in the 2x, 0, 2z Harker section of the simulated difference Patterson map. (a) shows the same section calculated from error-free data. Occupancy has been set here to one and the mean displacement of the cluster position to 3 Å.
Table 1  The Quality of the Data Sets Collected from Crystals of the Large Ribosomal Subunit from *Bacillus stearothermophilus* (B50S)*

<table>
<thead>
<tr>
<th>Data Set</th>
<th>r_{max} (Å)</th>
<th>r_{min} (Å)</th>
<th>Unique Ref.</th>
<th>R_{merge} (%)</th>
<th>Cell Dimensions (Å)</th>
<th>D(v) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>21</td>
<td>14.9</td>
<td>5561</td>
<td>6.9</td>
<td>297.6 547 383.3 111.6</td>
<td>+1.068</td>
</tr>
<tr>
<td>A</td>
<td>16</td>
<td>14.4</td>
<td>5226</td>
<td>7.8</td>
<td>302.9 547 374.4 112.2</td>
<td>+0.057</td>
</tr>
<tr>
<td>T</td>
<td>19</td>
<td>18.0</td>
<td>4891</td>
<td>7.2</td>
<td>300.0 547 377.0 111.9</td>
<td>0.000</td>
</tr>
<tr>
<td>D</td>
<td>48</td>
<td>14.0</td>
<td>4535</td>
<td>14.2</td>
<td>302.9 547 376.0 111.1</td>
<td>+0.618</td>
</tr>
<tr>
<td>E</td>
<td>41</td>
<td>12.8</td>
<td>8331</td>
<td>7.4</td>
<td>300.0 547 377.0 111.9</td>
<td>+0.057</td>
</tr>
<tr>
<td>R</td>
<td>53</td>
<td>15.7</td>
<td>4808</td>
<td>9.5</td>
<td>300.0 547 382.0 111.9</td>
<td>0.000</td>
</tr>
<tr>
<td>M</td>
<td>83</td>
<td>16.0</td>
<td>5104</td>
<td>10.4</td>
<td>293.1 547 372.8 111.4</td>
<td>-3.053</td>
</tr>
</tbody>
</table>

* a NAT: native; DER: derivative; M: mutant. D(v) = the change in crystal volume (compared to crystal 41).

The cell axes were scaled to a common b axis (d_{b0}=547 Å).

the map at the spherical angles (80, 60, 180). Thus, the threefold noncrystallographic axis is almost parallel to the crystallographic c axis (in the ac plane, 12° away from the c axis) The twofold axis is found in a plane containing the crystallographic b axis, perpendicular to the threefold axis, 30° away from the ac plane and 120° from the twofold along b and from its symmetry image.

Such a combination of noncrystallographic threefold and twofold axes suggests that there are three particles in the asymmetric unit. The self-rotation results can be related to electron micrographs of sections through three-dimensional crystals of B50S. According to these, one expects that a section parallel to the ab plane will show approximate noncrystallographic threefold symmetry in the asymmetric unit.

Self-rotation searches performed with the data collected from the crystals of mutated particles (83) showed an additional peak at $\kappa = 60^\circ$, suggesting 622 pseudo-symmetry. This result, together with the somewhat smaller unit cell obtained for this crystal, indicates that the packing in the crystal of the mutated particles may be different from that of the native and the derivatized crystals. Although this conclusion is based on the analysis of only one crystal of mutated particles, it is likely to be correct, as none of the other B50S crystals have shown this behavior.

**THE LOCATION OF THE GOLD CLUSTER BY DIFFERENCE PATTERSON SYNTHESIS**

Because of the observed lack of isomorphism, each data set of the native and the derivative was handled individually. Difference Patterson maps were calculated for every possible combination of native and derivatives at various resolution ranges. It was found that for obtaining consistent results, the resolution had to be limited, as including weak data at resolution higher than 20 Å increased the apparent lack of isomorphism and the experimental error, whereas the terms with resolution lower than 50 Å did not contribute to the phasing, as expected from the simulation studies.

In no case were we able to locate the goldcluster positions in the Harker sections by visual inspec-
Suitability of a Monojunctional Reagent

The only usable results evolved from the combination of data sets 46 (native) and 41 (derivative) (Table I). Their difference Patterson map showed highly correlated peaks, consistent with the existence of three undecagold clusters in the asymmetric unit (Figure 5), in accord with the results from the self-rotation searches and the motifs seen by electron microscopy in thin sections of the crystals (Figure 5). As in space group C2 only the $a$ and $z$ coordinates are fixed. Therefore the highest Patterson maximum was set to $y = 0$, and the $y$ coordinates of the other two positions were fixed relative to it.

### THE INITIAL SIR PHASE SET AND ITS REFINEMENT

The positions of the clusters, their occupancies and the overall temperature factor were refined in the resolution range 26–45 Å. The verified maxima are circled.

As seen, the solution of the structure of the 10S subunit is still at an early stage. However, the non-crystallographic threefold axis may be exploited to improve the phases by non-crystallographic averaging. The molecular replacement method may provide the base for these studies. This method is based on positioning a known model in the unit cell of the unknown structure. As this method was developed for studies at high resolution, and as in our case both the model, obtained by image reconstruction and the crystallographic data are of low resolution (20–28 Å), its suitability was tested. The calculated structure factors from a 50S model placed in a cubic box of a length of 225 Å, chosen to match the longest dimension of the particle, were used as input data. The model was filled with randomly and uniformly distributed scatterers and rotated by Euler angles, and cross-rotation searches at different resolution ranges and with different Patterson vector cutoffs, were run. This test was not complicated by the experimental error, allowed the choice of suitable resolution ranges and assessed the sensitivity of the cross-rotation results to the Patterson cutoff length and the size of the model cell.

Preliminary searches with the test data as well as the experimental structure factors were carried out using the program MERLOT with some modifications (Z. Otwinowski and J. Freedman, Yale University, and M. Eisenstein, Weizmann Institute). In the best runs, the difference between the height of the expected peak and the background (second peak) was taken as the criterion for the quality of the results. These tests showed that the meaningful results are obtained even at low resolu-
Table II  The Figure of Merit, Phase Difference and $R_{\text{value}}$ of the SIR Phase Sets Before and After Two Cycles of Solvent Flattening with Phase Combination

<table>
<thead>
<tr>
<th>Res. (Å)</th>
<th>SIR</th>
<th>SIR + 1 Cycle</th>
<th>SIR + 2 Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Res.</td>
<td>FOM</td>
<td>$\Delta \phi$</td>
</tr>
<tr>
<td>90.0–62.1</td>
<td>0.17</td>
<td>0.796</td>
<td>67</td>
</tr>
<tr>
<td>62.1–50.3</td>
<td>0.53</td>
<td>42</td>
<td>0.499</td>
</tr>
<tr>
<td>50.3–43.4</td>
<td>0.53</td>
<td>28</td>
<td>0.559</td>
</tr>
<tr>
<td>43.4–38.8</td>
<td>0.54</td>
<td>24</td>
<td>0.578</td>
</tr>
<tr>
<td>38.8–35.3</td>
<td>0.61</td>
<td>35</td>
<td>0.466</td>
</tr>
<tr>
<td>35.3–32.7</td>
<td>0.61</td>
<td>25</td>
<td>0.495</td>
</tr>
<tr>
<td>32.7–30.5</td>
<td>0.61</td>
<td>20</td>
<td>0.466</td>
</tr>
<tr>
<td>30.5–28.8</td>
<td>0.58</td>
<td>30</td>
<td>0.496</td>
</tr>
<tr>
<td>28.8–27.3</td>
<td>0.64</td>
<td>17</td>
<td>0.483</td>
</tr>
<tr>
<td>27.3–26.0</td>
<td>0.57</td>
<td>18</td>
<td>0.550</td>
</tr>
<tr>
<td>90.0–26.0</td>
<td>0.56</td>
<td>26</td>
<td>0.546</td>
</tr>
</tbody>
</table>

* FOM: figure of merit; $N_T$: number of reflections. $\Delta \phi$: the difference between the values of the phases before and after the phase combination (from the solvent-flattened map).

Inclusion of the very low resolution terms (below 80 Å) obscured the results, whereas the higher resolution terms led only to negligible improvement. Thus, the basic requirements for a successful application of the molecular replacement method can be fulfilled in our case. Furthermore, it may be possible to use the noncrystallographic symmetry to extend the phases to higher resolution.

In addition, the application of multichannel maximum entropy techniques aided by likelihood calculations$^{16,17}$ was predicted to yield considerable improvement in phases at higher resolution. The position of the gold clusters in conjunction with the structure factor differences between native and derivative crystals, as well as the noncrystallographic symmetry in conjunction with the existing solvent mask, should act as powerful constraints on the maximum entropy equations.

As has been described above, several lines of evidence support the validity of the current electron density map. Nevertheless, the rather low degree of isomorphism of the crystals of B50S directed our attention to an alternative phasing procedure, multiple-wavelength anomalous dispersion, which requires data collected from a single crystal at different wavelengths. For obtaining measurable anomalous signals, binding the gold or similar clusters in more than one site may be needed.

The suitability of the undecagold cluster for phasing data collected from crystals of the 50S subunits from B. stearothermophilus has encouraged us to use the undecagold cluster for phasing data collected from crystals of a much higher quality.

![FIGURE 6](image-url) A stereo plot of the whole unit cell, down the $z$ direction, of the electron density map after two cycles of solvent flattening and phase combination (resolution range 26–60 Å).
those of the large ribosomal subunits of *Halococcus marismortui*, which diffract to 2.9 Å. However, while attempting at the extension of the procedures developed for 50S to the halophilic ribosomes, we found that a straightforward approach was not possible, as there are no naturally exposed sulfhydrils on the surface of the halophilic large ribosomal subunit suitable for derivatization. Therefore site-directed mutagenesis for the insertions of exposed SH groups has been initiated.  

The studies presented here have been initiated under the inspiration and guidance of the late Prof. H. G. Wittmann and carried out with enthusiasm together with the late Dr. K. von Bohlen. We would like to warmly thank our colleagues from Hamburg, Berlin, and Rehovot for their active participation and illuminating discussions. We also want to thank Dr. W. Jahn for his efforts in the development and the production of the undecagold cluster and its monofunctional reagent, and Drs. T. Leighton and J. Schnier for producing the mutant.

Data were collected at the following synchrotron facilities: EMBL and MPG beam lines at DESY, Hamburg; CHESS, Cornell University; SSRL, Stanford University; and PF/KEK, Japan.

Support was provided by the Max-Planck Society, the National Institute of Health (NIH GM 34360) and the Kimmelman Center for Macromolecular Assembly at the Weizmann Institute. AY holds the Martin S. Kimmel Professorial Chair.

REFERENCES