Small-angle X-ray Scattering Studies of the Iron-Molybdenum Cofactor from Azotobacter vinelandii Nitrogenase*

(Received for publication, December 31, 1992, and in revised form, April 22, 1993)

David Eliezer§, Patrick Frank‡, Nanna Gillis§, William E. Newton¶, Sebastian Doniach§, and Keith O. Hodgson§**††

From the §Department of Physics, the ‡Department of Chemistry and ††Stanford Synchrotron Radiation Laboratory, and the ¶Department of Applied Physics, Stanford University, Stanford, California 94305 and the §§Department of Biochemistry and Nutrition, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

The nitrogenase enzyme complex, consisting of the molybdenum-iron protein and the iron protein, plays a critical role in the biological reduction of dinitrogen to ammonia (nitrogen fixation). The nitrogen-fixing site within the molybdenum-iron protein is an iron-molybdenum-sulfur cofactor (FeMoco) of roughly 1000-2000 Dalton mass. Structural aspects of FeMoco have been determined by spectroscopic and more recently by crystallographic studies. In order to determine the radius of gyration (Rg) of isolated FeMoco, we have performed small-angle x-ray scattering studies of FeMoco in N-methylformamide solution, in the absence of the molybdenum-iron protein. Model compounds of known structure have also been examined in similar solvents, N,N-dimethylformamide and acetonitrile, as controls and for calibration purposes. The Rg values obtained for the models are in good agreement with calculations based upon their respective crystal structures. However, the Rg obtained for FeMoco clearly indicates that the cofactor is not monomeric in solution, but rather aggregated and possibly polydisperse. Further, Rg values were also measured after addition of thiol, dithionite, and thiol and dithionite, to the FeMoco samples. The results indicate, surprisingly, that oxidation state and putative thiol coordination have no detectable effect on the aggregation behavior of FeMoco in solution, as determined by these measurements.

Biological reduction of dinitrogen is catalyzed by the nitrogenase molybdenum-iron protein includes a proposed inorganic cluster model to probe the size and shape of the iron-molybdenum-sulfur cofactor (FeMoco) of roughly 3 pm3 radius of gyration. The recent report of a preliminary crystal structure for the nitrogenase molybdenum-iron protein includes a proposed

3. FeMoco has an elemental composition of Mo:Fe:S:N:O=1:1:2:1:1. Its homocitrate (4) and can be extruded in a native-like state into polar solvents from the holoenzyme. It is also capable of restoring full spectroscopic and catalytic activity to cofactor-deficient molybdenum-iron proteins produced in a variety of mutant organisms (1, 5, 6). FeMoco in N-methylformamide solution exhibits electrochemical behavior analogous to that observed within the native enzyme (7, 8). This includes the three redox states FeMoco(ox), FeMoco(s-r), and FeMoco(red), each separated by single electron transfers. For recent reviews, see Refs. 5 and 9.

Until recently, the primary insights into the metrical details of FeMoco have come from x-ray absorption spectroscopy studies (edge and extended fine structure) (10–12), which clearly demonstrated the involvement of molybdenum in a polynuclear cluster bridged through a first shell of sulfides to a second shell of iron, and also provided evidence for larger metal–metal interactions. Investigations utilizing EPR have suggested a single binding site for cysteine-thiol (5), and a site for histidine-imidazole binding has been inferred from electron nuclear double resonance and electron spin echo envelope modulation EPR measurements on altered molybdenum-iron proteins produced by directed mutagenesis (3, 13–15). In addition, evidence for solution heterogeneity has been obtained from electrochemical and other studies (16, 17). The recent report of a preliminary crystal structure for the nitrogenase molybdenum-iron protein includes a proposed structural model of the FeMoco cluster bound within its protein matrix, although metrical details are not available at the published level of resolution and refinement (4).

The availability of high-flux synchrotron x-ray sources makes feasible the extension of the solution small-angle x-ray scattering (SAXS) technique to experiments involving small or weakly scattering molecules. Here we report the results of our experiments using synchrotron radiation small-angle x-ray scattering from solutions of FeMoco and related inorganic cluster models to probe the size and shape of FeMoco in solution. Experiments have been carried out on two of the three biologically relevant oxidation states of FeMoco, FeMoco(ox) and FeMoco(s-r), both alone and in the presence of added benzenethiol.

EXPERIMENTAL PROCEDURES

The preparation and activity verification of the FeMoco from Azotobacter vinelandii nitrogenase were carried out as described elsewhere (18, 19). All reagents were rendered oxygen-free using standard Schlenk methods. N-Methylformamide was purified by stirring overnight over sodium bicarbonate, followed by vacuum distillation from solid barium oxide. N,N-Dimethylformamide and acetonitrile were distilled, the former under vacuum, from calcium hydride. All distilled solvents were further degassed under vacuum and stored under a...
purified dinitrogen atmosphere. All the following preparations were performed in a Vacuum-Atmospheres glove box operating under a dinitrogen atmosphere of 1 part/million dioxygen or less.

The crude, N-methylformamide solutions of FeMoco as extracted from the molybdenum-iron protein were further purified by anaerobic gel filtration under vacuum. SAXS experiments the FeMoco(ox) samples were used as obtained following the concentration step. Final sample concentrations ranged from 1 to 3 mM FeMoco. FeMoco(ox)- samples were prepared using a 5 × stoichiometric excess of sodium dithionite, added as a 0.2 M solution in 1 mM NaOH. Benzenethiol, added as needed in 5 × stoichiometric excess, was used either as a 1% (97 mM) solution, or as a 250 mM solution, in N-methylformamide. The samples were frozen under dry ice in crimp top or screw-cap v-bottom vials under a silicone/Teflon septum, covered with at least four layers of Saran Wrap, and they will still use (3-4 days later).

Immediately following the SAXS experiments, the sealed capillaries (see below) containing FeMoco were frozen in dry ice. The frozen samples were later (4 days) recovered from the capillaries and converted into samples for EPR and activity analysis. The EPR samples were stored in liquid nitrogen inside 1-mm inner diameter serum-capped quartz tubes. The activity samples were treated with a large excess of sodium dithionite, adding small volumes as a 1 M solution in 10 mM sodium hydroxide, and again stored in crimp-top or screw-cap v-bottomed vials under a silica/Teflon septum, covered with at least four layers of Saran Wrap. All but two samples (see below) gave good EPR activities.

The inorganic complexes (C3H7)NCH2C6H5][MoFe3S6(SC2H5)12] and ((C3H7)NCH2C6H5)[MoFe3S6(SC2H5)12] were prepared by literature methods (20, 21), and their identity and purity were established by elemental analyses and electronic spectra. The compound (C3H7)NCH2C6H5][MoFe3S6(SC2H5)12] was a kind gift of Professor Richard H. Holm of the Department of Chemistry, Harvard University. Sample solutions of the molybdenum-containing dicubanes were prepared in N,N-dimethylformamide by dilution of a more concentrated stock solution. The sample solutions of the [(NMe2)2Si][Mo2Fe6S8(SC2H5)12] complex were prepared in similar fashion in acetonitrile. Final sample concentrations ranged from 1 to 10 mM. All model solutions were sealed in 1-ml crimp-top vials and stored in dry ice prior to use.

Following the scattering experiments, sample integrity of the model compounds was examined both by means of UV/visible spectroscopy and by examining the availability, or unavailability, of cluster iron to capture by o-phenanthroline. Several model compound solutions was examined both by means of UV/visible spectroscopy for all samples, including the model compounds. However, a fitting range was chosen so as to be within the Guinier region, for an Rg value of approximately 5 Å, but as much as possible outside the very small angle regions where the observed aggregation could introduce systematic errors. Within this range, the data were observed to be satisfactorily linear, as required by Guinier's theory. Fits were performed using a linear least squares fitting procedure, and representative fits are shown in Fig. 1. Averaged fit results are summarized in Table I. Statistical errors in the data were correctly propagated through the fitting procedure and through the subsequent averaging steps. The resulting error estimates are included in Table I. No significant dependence of the fit results on the different FeMoco and model compound concentrations used was observed, indicating that interparticle interference was not a problem. The observed aggregation is discussed below.

As first shown by Guinier, solution SAXS data from monodisperse samples is well approximated by a Gaussian shape in the Guinier range, defined by 2πRg/S < 1 (26, 27). The width of the Gaussian can be obtained from a straight line fit to a (Guinier) plot of log I versus S² (where I is the scattered intensity), and is proportional to Rg. However, aggregation and, in the case where sample concentrations are high, interparticle interference effects can cause data to deviate from a Gaussian shape and can skew estimates of Rg obtained from Guinier plots. These effects are most pronounced at very small angles.

When the treated SAXS data were represented in Guinier plots, evidence of some degree of aggregation was observed for all samples, including the model compounds. However, a fitting range was chosen so as to be within the Guinier region, for an Rg value of approximately 5 Å, but as much as possible outside the very small angle regions where the observed aggregation could introduce systematic errors. Within this range, the data were observed to be satisfactorily linear, as required by Guinier's theory. Fits were performed using a linear least squares fitting procedure, and representative fits are shown in Fig. 1. Averaged fit results are summarized in Table I. Statistical errors in the data were correctly propagated through the fitting procedure and through the subsequent averaging steps. The resulting error estimates are included in Table I. No significant dependence of the fit results on the different FeMoco and model compound concentrations used was observed, indicating that interparticle interference was not a problem. The observed aggregation is discussed below.

The model inorganic clusters [Mo2Fe3S6(SC2H5)12]⁻ and [Mo2Fe3S6(SC2H5)12]⁺ were chosen for their similarity to FeMoco in composition, size range, and air sensitivity and were in fact originally synthesized as possible models for the

---

**RESULTS**

As first shown by Guinier, solution SAXS data from monodisperse samples is well approximated by a Gaussian shape in the Guinier range, defined by 2πRg/S < 1 (26, 27). The width of the Gaussian can be obtained from a straight line fit to a (Guinier) plot of log I versus S² (where I is the scattered intensity), and is proportional to Rg. However, aggregation and, in the case where sample concentrations are high, interparticle interference effects can cause data to deviate from a Gaussian shape and can skew estimates of Rg obtained from Guinier plots. These effects are most pronounced at very small angles.

When the treated SAXS data were represented in Guinier plots, evidence of some degree of aggregation was observed for all samples, including the model compounds. However, a fitting range was chosen so as to be within the Guinier region, for an Rg value of approximately 5 Å, but as much as possible outside the very small angle regions where the observed aggregation could introduce systematic errors. Within this range, the data were observed to be satisfactorily linear, as required by Guinier's theory. Fits were performed using a linear least squares fitting procedure, and representative fits are shown in Fig. 1. Averaged fit results are summarized in Table I. Statistical errors in the data were correctly propagated through the fitting procedure and through the subsequent averaging steps. The resulting error estimates are included in Table I. No significant dependence of the fit results on the different FeMoco and model compound concentrations used was observed, indicating that interparticle interference was not a problem. The observed aggregation is discussed below.

The model inorganic clusters [Mo2Fe3S6(SC2H5)12]⁻ and [Mo2Fe3S6(SC2H5)12]⁺ were chosen for their similarity to FeMoco in composition, size range, and air sensitivity and were in fact originally synthesized as possible models for the
structure of FeMoco (20, 21). The values of $R_g$ obtained by direct calculation from the atomic coordinates of the six Fe and seven Fe compounds are 4.1 and 5.1 Å, respectively. The values of $R_g$ obtained from the data are 4.22 and 4.53 Å, respectively, in good agreement. The model compound $[\text{Na}_2\text{Fe}_{16}\text{S}_{20}]^{10-}$ was chosen for its different shape and size. The $R_g$ value obtained by calculating moments of inertia from the molecular dimensions provided in the publication of the structure (28) is 6.4 Å, and the measured experimental value is 6.77 Å, again in good agreement. The good agreement between the experimental $R_g$ values obtained for the model compounds and those calculated from the respective crystal structures is a good indication that the Guinier analysis used to extract the $R_g$ is valid, and that the effects of aggregation in these samples do not extend into the fitting range used. We believe that the aggregation detected in these samples consists of only trace amounts of large aggregates. In all of the instances measured, the data for FeMoco(ox) and FeMoco(s-ri), both as such and with added thiol, were nearly identical to each other and gave very similar values of $R_g$. As a result these values were averaged together, and the value reported in Table I is this average. The final $R_g$ value of 6.96 Å obtained for FeMoco is surprisingly high, considering that its molecular weight, as calculated from its minimal composition, is significantly lower than that of any of the models involved, and that its overall shape in the protein (4) is similar to that of the dicubane model compounds. In addition, the FeMoco SAXS data showed significantly stronger aggregation effects than the model compound data (Fig. 2). These two observations lead us to hypothesize that unlike in the case of the model compounds, the aggregation effects observed in the FeMoco data are not caused merely by trace amounts of large aggregates. Instead, we believe that a distribution of smaller FeMoco oligomers is most likely responsible for the observed $R_g$ and for the large deviation of the data from a Gaussian shape in the smaller angle regions. The following discussion supports this hypothesis.

The intensity scattered by various components in a solution increases as the square of the mass of the scatterer, but falls off much more sharply with angle as the scatterer increases in size. Hence, large aggregates in solution samples invariably produce more scattering in the very low angle regions than monomers do. At higher angles, however, the scatter is dominated by smaller components (29). We therefore expect that the effect of the larger aggregates in our samples is negligible in the relatively higher angle region used for the Guinier analysis. This is verified by the model compound results, where in spite of the presence of trace amounts of large aggregates, Guinier analysis yields $R_g$ figures in good agreement with calculations. We argue qualitatively then that relatively small aggregates are necessary to explain the larger than expected $R_g$ value measured for FeMoco.

In order to pursue this reasoning in a more quantitative manner, we applied a model using the analytical expression for the scattering from homogeneous spheres of specified $R_g$, first derived by Rayleigh (27, 30).

$$I(S) = \left[\sin(2\pi SR) - 2\pi SR \cos(2\pi SR)/(2\pi SR)^2\right]^2$$ (Eq. 1)

where $R$ is the sphere radius. Using this formula, the theoretical scattering pattern for a sample consisting of spheres of any single size can be calculated. This pattern can be scaled so that it intersects the measured FeMoco SAXS data at one

![Fig. 1. Normalized background-subtracted SAXS data from four different FeMoco samples represented in Guinier plots (log $I$ versus $S^2$) along with the fits from which $R_g$ values are obtained. The fitting region is well inside the region of reliable data, which was determined in this case to start at $S = 0.012$ ($S^2 = 0.00015$). The plots have been shifted arbitrarily along the ordinate for clarity of presentation. The thiol-containing samples were prepared directly from the final thiol-free sample solutions. No simple correlation was found between FeMoco concentration and FeMoco(s-ri), both as such and with added thiol, were nearly identical to each other and gave very similar values of $R_g$. An average of two samples.

An average of six samples.

An average of three samples.

An average of two samples.

An average of two samples.](image)

![Fig. 2. Scattering data from equal concentrations of FeMoco and the $[\text{Mo}_2\text{Fe}_6\text{S}_4(\text{SC}_5\text{H}_4)_2]^{1-}$ model compound at angles smaller than those used for the Guinier fits. The sharp rise of the scattering at small angles indicates the presence of trace amounts of large aggregates. The presence of more and/or larger aggregates in the FeMoco sample is readily apparent from the much stronger scattering at these low angles.](image)
point, but does not exceed it at any point. It will then represent the maximum contribution that the presence, in the sample, of homogeneous spheres of that size could make to the measured data. Although the FeMoco aggregates are not known to be exactly spherical or entirely homogeneous, we believe that this model is still useful for gauging the contribution of aggregates to the data.

Using successively larger spheres, it is straightforward to estimate the maximum contribution that each particle size could make to the data (Fig. 3). The contribution to the Guinier fitting region can then be calculated as the integrated intensity from the model normalized by the integrated intensity from the data (Fig. 4). It is clear that according to this model the contribution from aggregates to the FeMoco data in the fitting region is limited to particles of less than about 6 25-Å $R_g$. In fact Fig. 4 suggests that the major contribution to the data comes from particles of a 6 to $\sim$12-Å $R_g$. This bears out our expectation that the presence of smaller aggregates is necessary to explain the measured FeMoco $R_g$ value.

Another parameter which can be derived from SANS data is the forward scattered intensity ($I_0$). $I_0$ is proportional to $c(\rho_s - \rho_b)V^2$ (where c is the concentration of the scattering molecule, V its volume, $\rho_s$ its electron density, and $\rho_b$ is the electron density of the background solvent) (27). The ratio of $I_0$ for two monodisperse solutes of similar electron densities in identical environments approximately equals the square of the ratio of their masses. Thus measurements of $I_0$ can be used to determine the mass of a solute. It is important to note, however, that $I_0$ is not independent of the aggregation state of the solute because $I_0$ for an n-mer is proportional to $(nV)^2$ whereas $I_0$ for n monomers is proportional to $nV^2$.

Because of this fact, using $I_0$ to obtain a meaningful measure of the mass of solutes which aggregate is difficult.

Although the aggregation present in our samples makes a precise analysis using $I_0$ impracticable, our Guinier fits do provide extrapolated values of $I_0$. In the case of the model compounds, these values should be representative of the true monomeric $I_0$ because the data in the fitting range was seen to be reasonably free of contributions from aggregates. Unfortunately, $I_0$ values for the FeMoco samples represent only some weighted average of the $I_0$ values of the various oligomers contributing to the scattering in the fitting range. Nevertheless in Table II we list the square root of the ratio of the concentration-normalized $I_0$ for each of the model compounds to that of FeMoco. This is a measure of the ratio of the apparent masses. We also list the value of this ratio that would be expected for monomeric samples. Contrary to the expected values, the apparent mass of the FeMoco samples is observed to be greater than that of the model compounds. This provides further evidence that the FeMoco samples are not monomeric. In fact the apparent mass of FeMoco is about 2.5 to 3 times greater than expected. This in turn supports the conclusion that a significant portion of the FeMoco aggregates are rather small. We note that the measured ratios of the masses of the various models to each other are in good agreement with the expected values. While this $I_0$ analysis should not be over-interpreted, it does support the conclusions we have drawn from our previous analyses.

In summary, the model compound data can be well explained by assuming a monomeric sample with trace amounts

![Fig. 3. Maximum possible contributions from homogeneous spheres of varying $R_g$ to the measured intensity. The scattering calculated for each sphere size may not exceed the measured scattering at any point.](image)

![Fig. 4. Maximum percentage contribution from homogeneous spheres of varying $R_g$ to the FeMoco(ox) scattering data in the range used for the Guinier fits. Calculated as the integral, over the fitting range, of the maximum possible sphere model intensity divided by the integral, over the same range, of the measured intensity.](image)

<table>
<thead>
<tr>
<th>Sample</th>
<th>$M/M(\text{FeMoco})$ observed</th>
<th>$M/M(\text{FeMoco})$ expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>FeMoco</td>
<td>$1.00 \pm 0.03^a$</td>
<td>$1.00^b$</td>
</tr>
<tr>
<td>$[\text{Mo}_6\text{Fe}_3\text{S}_4(\text{SC}_2\text{H}_5)_2]^+$</td>
<td>$0.70 \pm 0.04^a$</td>
<td>$1.80^b$</td>
</tr>
<tr>
<td>$[\text{Mo}_6\text{Fe}_3\text{S}_4(\text{SC}_2\text{H}_5)_2]^+$</td>
<td>$0.54 \pm 0.02^a$</td>
<td>$1.55^b$</td>
</tr>
<tr>
<td>$[\text{Na}_3\text{Fe}_6\text{S}_6]^-$</td>
<td>$1.06 \pm 0.03^a$</td>
<td>$2.90^b$</td>
</tr>
</tbody>
</table>

$^a$ An average of six samples.

$^b$ Calculated from the molecular formula for FeMoco in (4) using molybdenum, iron and sulfur atoms.

$^c$ An average of three samples.

$^d$ Calculated from the molecular formula using molybdenum, iron, sodium, and sulfur atoms.

$^e$ An average of two samples.
of large aggregates. In contrast, the large measured value of $R_g$ for FeMoco cannot be explained in the same way. Our simple modeling analysis suggests that a significant presence of small FeMoco oligomers is necessary to explain the data. A unique distribution of particle sizes cannot be obtained from the data. However, we conclude with some confidence that in N-methylformamide solution, FeMoco largely exists in aggregated states of up to 25 Å in $R_g$. The average $R_g$ value of the population is approximately 7 Å, a value significantly larger than what would be expected for a monomer. In data are difficult to interpret for aggregate-containing samples, but do nevertheless support these conclusions.

DISCUSSION

Although the determination of $R_g$ for monomeric FeMoco in solution did not prove feasible in these experiments, the results we have obtained are interesting for a number of reasons. The aggregation of FeMoco in solution may explain the failure of persistent widespread efforts to produce crystalline FeMoco samples suitable for crystallographic studies. It may also explain the noted slow electron exchange between FeMoco(ox) and FeMoco(s-r) in solution (5), as well as the apparent existence in solution of multiple FeMoco species with different electrochemical properties (7, 16). The shielding of interior components of FeMoco aggregates against an electrode surface may explain the sluggish or incomplete electroactivity exhibited by FeMoco solutions at carbon electrodes (7). On the other hand these aggregates must be porous enough to admit small molecules such as benzene-thiol because the stoichiometry of thiol interaction with FeMoco under the above conditions is thought to be 1:1 (32). It is also possible that the aggregate distribution of FeMoco in solution is not always the same. This may explain the batch- and history-related variation of FeMoco UV/visible spectra (17) and the variation with pH of FeMoco EPR spectra (5, 16).

It should be possible in the future to use SAXS to screen FeMoco samples in order to search for conditions under which aggregation does not occur. In fact our studies of different oxidation and thiol ligation states of FeMoco are the beginning of such a search. However, the lack of any observed variation in the SAXS results for these samples suggests that previously observed effects on FeMoco(s-r) EPR spectra due to thiol ligation (5) are not a result of some modification of aggregation behavior in solution.

There is a considerable interest in the solution chemistry of FeMoco as an inorganic heterometallic cluster. With the emergence of a FeMoco structure within the molybdenum-iron protein, for example, direct laboratory synthesis has become an immediate goal. The conditions necessary to induce catalytic activity in isolated FeMoco are also of great interest. Knowledge of the solution behavior of FeMoco is of vital importance in approaching such issues.

Acknowledgments—The critical interest and technical help of Dr. Michel Koch, Dr. Hirotugu Tsuruta, and Dr. Brit Hedman are gratefully acknowledged.

REFERENCES