The folding and assembly of the dodecameric type II dehydroquinases

Nicholas C. PRICE^{*1}, Deborah J. BOAM^{*}, Sharon M. KELLY^{*}, Doris DUNCAN^{*}, Tino KRELL[†], David G. GOURLEY[‡], John R. COGGINS[‡], Richard VIRDEN[‡] and Alastair R. HAWKINS[‡]

^{*}Department of Biological Sciences, University of Stirling, Stirling FK9 4LA, Scotland, U.K., [†]Division of Biochemistry and Molecular Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, Scotland, U.K., and [‡]Department of Biochemistry and Genetics, Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, U.K.

The dodecameric type II dehydroquinases (DHQases) have an unusual quaternary structure in which four trimeric units are arranged with cubic 23 symmetry. The unfolding and refolding behaviour of the enzymes from *Streptomyces coelicolor* and *Mycobacterium tuberculosis* have been studied. Gel-permeation studies show that, at low concentrations (0.5 M) of guanidinium chloride (GdmCl), both enzymes dissociate into trimeric units, with little or no change in the secondary or tertiary structure and with a 15 % loss (*S. coelicolor*) or a 55 % increase (*M. tuberculosis*) in activity. At higher concentrations of GdmCl, both enzymes undergo sharp unfolding transitions over narrow ranges of the denaturant concentration, consistent with co-operative unfolding of the subunits. When the concentration of GdmCl is lowered by dilution from 6 M to 0.55 M, the enzyme from *S. coelicolor* refolds in an efficient manner to form trimeric units, with more than 75 % regain of activity. Using a similar approach the *M.*

tuberculosis enzyme regains less than 35 % activity. From the time courses of the changes in CD, fluorescence and activity of the *S. coelicolor* enzyme, an outline model for the refolding of the enzyme has been proposed. The model involves a rapid refolding event in which approximately half the secondary structure is regained. A slower folding process follows within the monomer, resulting in acquisition of the full secondary structure. The major changes in fluorescence occur in a second-order process which involves the association of two folded monomers. Regain of activity is dependent on a further associative event, showing that the minimum active unit must be at least trimeric. Reassembly of the dodecameric *S. coelicolor* enzyme and essentially complete regain of activity can be accomplished if the denatured enzyme is dialysed extensively to remove GdmCl. These results are discussed in terms of the recently solved X-ray structures of type II DHQases from these sources.

INTRODUCTION

The study of the refolding of unfolded proteins has provided a great deal of valuable information on the likely pathways by which proteins fold *in vivo* [1–4]. Using a number of small model proteins, the relationships between acquisition of native secondary and tertiary structures and the regain of biological activity have been explored in detail. The involvement of compact intermediate states, often termed ‘molten globules’, has been shown to be important in a number of cases such as β -lactoglobulin [5] and the β_2 unit of tryptophan synthase [6], whereas in others, such as the acyl-CoA-binding protein [7], there appears to be little kinetic evidence for the participation of discrete intermediates. In the cases of lysozyme [8] and the *trp* aporepressor [9], a number of parallel pathways of folding are involved. Thus it may be misleading to imagine that there is a single universally applicable model of protein folding.

The refolding of oligomeric enzymes raises a number of additional problems, of which the most important relates to the specificity of association of subunits. As shown in a number of studies, the association of subunits is very specific so that the recognition sites on a subunit must be formed during the folding of that subunit. However, the folded structure of the isolated subunit can have only very marginal stability in the absence of the other subunit(s), since the interaction site is largely hydrophobic in nature [1,10]. Folding and assembly *in vivo* have been shown to involve the participation of a number of proteins including molecular chaperones, as well as enzymes such as

peptidylprolyl isomerase and protein disulphide-isomerase [2,4,11].

As part of a programme of investigation of the folding and assembly of oligomeric enzymes, we have studied the unfolding and refolding of the type II dehydroquinases (DHQases) from *Streptomyces coelicolor* and *Mycobacterium tuberculosis*. The type II DHQases can play either a biosynthetic role in the shikimate pathway or a catabolic role in the quinate pathway and are known to catalyse an *anti* elimination reaction, possibly proceeding via an enolate intermediate pathway [12]. These enzymes have a number of interesting features for this type of study. First, they are complex structures each consisting of twelve subunits of molecular mass around 16 kDa. The refolding of such structures could occur by a number of possible assembly mechanisms. Secondly, the high-resolution structures of the *M. tuberculosis* (D. G. Gourley, A. K. Shrive, I. Polikarpov, T. Krell, A. R. Hawkins, L. Sawyer, N. W. Isaacs and J. R. Coggins, unpublished work) and *S. coelicolor* (A. Roszak, D. G. Gourley, T. Krell, J. R. Coggins and A. J. Laphorn, unpublished work) enzymes have been recently solved to 0.21 nm (2.1 Å) and 0.18 nm (1.8 Å) resolution respectively. The twelve subunits of the enzyme are packed in an unusual fashion as a tetramer of trimeric units with cubic 23 symmetry (an arrangement which has also been observed in the catabolic form of ornithine transcarbamoylase from *Pseudomonas aeruginosa* [13,14] and *Pyrococcus furiosus* [15]). Each subunit of the type II DHQases has a $\beta\alpha$ structure similar to that of flavodoxin [16], consisting of a five-stranded parallel β -sheet core flanked by four α -helices. The connections

Abbreviations used: DHQases, dehydroquinases; GdmCl, guanidinium chloride.

¹ To whom correspondence should be addressed (e-mail ncp1@stirling.ac.uk).

of the successive $\beta\alpha$ units follow the standard right-handed rule, which means that helices 2 and 3 are on one side of the sheet and helices 1 and 4 are on the other. Thirdly, both enzymes have a single tryptophan residue per subunit (located in the α -helix between strands 2 and 3 of β -sheet), the fluorescence of which provides a convenient measure of changes in the exposure of this residue as a result of folding within a subunit or interactions between subunits.

EXPERIMENTAL

The type II DHQases from *S. coelicolor* and *M. tuberculosis* were prepared as described previously [17,18]. Concentrations of enzymes were determined spectrophotometrically at 280 nm using absorption coefficients of 0.68 and 0.60 for 1 mg/ml solutions respectively. Assays of enzyme activity were performed spectrophotometrically at 234 nm [19]. The concentration of ammonium dehydroquinone used routinely was 1 mM. Under these conditions the specific activities of the enzymes from *S. coelicolor* and *M. tuberculosis* were 1710 and 36 $\mu\text{mol}/\text{min}$ per mg respectively.

Guanidinium chloride (GdmCl; Ultrapure grade) was purchased from Gibco-BRL Life Technologies. The concentrations of solutions of GdmCl were checked by refractive-index measurements [20].

Fluorescence measurements were made at 20 °C using a Perkin-Elmer LS50B spectrofluorimeter. CD measurements were made at 20 °C using a JASCO J-600 spectropolarimeter.

Stopped-flow measurements of the refolding process were made on an Applied Photophysics SX-17MV stopped-flow reaction analyser with a 10:1 mixing ratio. In the CD mode the instrument used an umbilical cell which involves a dead time estimated as 8 ms by using the well-characterized refolding of lysozyme after denaturation [8,21]; in the fluorescence mode, the dead time estimated using the reaction between 2,6-dichlorophenol-indophenol and L-ascorbic acid [22] is 1.7 ms.

Unfolding of the enzymes by GdmCl was monitored after incubation in the stated concentration of denaturing agent for 24 h in 50 mM Tris/HCl buffer, pH 7.5, containing 1 mM dithiothreitol. The catalytic activity of samples was determined in the presence of the same concentration of GdmCl as had been used in the denaturing procedure. The refolding process was initiated by dilution of enzymes that had been incubated in 6 M GdmCl with 10 vol. of buffer so as to give a final concentration of 0.55 M GdmCl. The concentration of GdmCl used for unfolding was chosen to ensure that the enzyme was fully unfolded (see below) and the extent of dilution was imposed by the stopped-flow instrument in which a 10:1 mixing ratio is the largest that is practicable. The refolding of the enzymes was monitored by changes in spectroscopic properties (CD and fluorescence) and the regain of catalytic activity. In all cases, control samples which were incubated in the final concentration of GdmCl (0.55 M) were included.

Gel-permeation studies were performed on a Superose 12 HR10/30 gel-permeation column at a flow rate of 0.2 ml/min, with detection by a Shimadzu SPD-M10AV diode-array detector. Typically 0.25 ml of a solution of protein (80 $\mu\text{g}/\text{ml}$) was applied to the column, with detection at 214 nm and 280 nm. Calibration of the column was performed using the following proteins as molecular-mass markers with the corresponding elution volumes shown: apoferritin (443 kDa; 8.48 ml), β -amylase (200 kDa; 8.88 ml), yeast alcohol dehydrogenase (150 kDa; 9.80 ml), BSA (66 kDa; 10.44 ml), carbonic anhydrase (29 kDa; 12.80 ml) and cytochrome *c* (12.4 kDa; 14.60 ml).

RESULTS

Unfolding of type II DHQases by GdmCl

In the absence of GdmCl, the type II DHQases from both *S. coelicolor* and *M. tuberculosis* were eluted from the gel-permeation column at a volume (8.95 ml) corresponding to an apparent molecular mass of 190 kDa (equal to the known mass of the dodecameric unit [12]). After a 24 h incubation in the presence of 0.55 M GdmCl, the elution volume (11.05 ml) of both enzymes corresponded to a molecular mass of 50 kDa, i.e. a globular trimeric unit. Preliminary studies showed that the complete dissociation of the dodecamer to the trimers occurred at concentrations of GdmCl as low as 0.2 M. The elution volumes of marker proteins (BSA and carbonic anhydrase) were not altered at these concentrations of GdmCl. It is noteworthy that the native dodecameric structure of ornithine transcarbamoylase from *P. aeruginosa* is much more stable towards denaturants; dissociation of the wild-type enzyme occurred on incubation with 5 M urea to yield inactive trimeric units [13]. By introducing a number of mutations, especially in the C-terminal region, dissociation of the *P. aeruginosa* enzyme was observed in the absence of denaturants [14].

The changes in far-UV CD, fluorescence and activity of both type II DHQases on incubation with GdmCl for 24 h were monitored. In each case, incubation in 6 M GdmCl leads to essentially complete loss of structure, as indicated by (i) the change in the far-UV CD signal, with at least a 10-fold decrease in ellipticity at 225 nm, and (ii) the shift of the fluorescence emission maximum from 340 nm (*S. coelicolor*) or 338 nm (*M. tuberculosis*) to 356 nm. The changes in CD (ellipticity at 225 nm), fluorescence intensity (at 350 nm) and catalytic activity are shown in Figure 1 as a percentage of the total change occurring between 0 and 6 M GdmCl. (It should be noted that, in the case of the *S. coelicolor* enzyme, the unfolding by 6 M GdmCl was accompanied by a 3-fold decrease in the intensity of the fluorescence at 350 nm, whereas in the case of the *M. tuberculosis* enzyme there was a 1.6-fold increase in fluorescence at this wavelength.)

For the *S. coelicolor* enzyme the changes in fluorescence and CD occur roughly in parallel with a sharp transition centred at about 3.75 M GdmCl. By contrast the activity is lost at much lower concentrations of GdmCl, with 15 % activity being lost at 0.5 M and 50 % at 1 M GdmCl.

For the *M. tuberculosis* enzyme, the midpoint for the CD and fluorescence changes was at about 4.25 M GdmCl, with again a sharp transition noted. In this case, however, the enzyme showed significant activation by low concentrations of GdmCl; this may reflect an increase in the flexibility of the enzyme, which would increase the rate of slow structural changes in the enzyme during the catalytic cycle (D. J. Boam, S. M. Kelly and N. C. Price, unpublished work).

The sharp nature of the changes in CD and fluorescence suggests that the overall secondary and tertiary structures of the enzymes are lost in a process in which the subunits unfold in a highly co-operative fashion. From the data available, there is little if any evidence for the existence of significantly populated intermediates with non-native secondary or tertiary structures.

Refolding of type II DHQases

Changes in activity during refolding

In the light of the above data, it is clear that incubation in 6 M GdmCl is required to bring about complete unfolding of the enzymes. Although the rate of the unfolding process was not investigated in detail, it was found that incubation in 6 M

Figure 1 Unfolding of type II DHQase by GdmCl

(a) Data for the *S. coelicolor* enzyme; (b) data for the *M. tuberculosis* enzyme. In both cases, enzyme (0.4 mg/ml) was incubated with the stated concentration of GdmCl in 50 mM Tris/HCl buffer, pH 7.5, containing 1 mM dithiothreitol for 24 h at 20 °C, prior to measurements of the various parameters. Squares, circles and triangles represent the changes in activity, ellipticity at 225 nm and fluorescence at 350 nm respectively. The data have been normalized to refer to the total change occurring between 0 and 6 M GdmCl as 100%; the negative values for activity at low [GdmCl] shown in (b) thus represent activation of the enzyme.

GdmCl for periods ranging from 1 to 24 h led to complete loss of structure of both enzymes, as indicated by changes in fluorescence, far-UV CD and activity. In practice, a 4 h incubation in 6 M GdmCl was used routinely prior to the initiation of refolding.

Refolding was initiated by dilution of the denatured enzymes with 10 vol. of buffer, so as to lower the concentration of GdmCl to 0.55 M. When the results from the gel-permeation studies are considered, it is clear that this strategy allows attention to be focused on the steps involved in refolding and assembly of the trimeric units from the unfolded monomers.

Most of the activity was regained in the case of the *S. coelicolor* enzyme, but this was not the case for the *M. tuberculosis* enzyme. Thus refolding of the enzyme from *S. coelicolor* at a concentration of 80 $\mu\text{g/ml}$ led to a regain of more than 75 % activity compared with a control sample which had been incubated with 0.55 M GdmCl (the residual concentration after 10:1 dilution). The extent of regain of activity was found to depend on protein concentration, falling to 56 % at 25 $\mu\text{g/ml}$ and 44 % at 12.5 $\mu\text{g/ml}$.

Figure 2 Kinetics of regain of activity of the type II DHQase from *S. coelicolor* during refolding

Following denaturation, refolding was initiated by dilution of the GdmCl from 6 M to 0.55 M. Curves A, B and C represent enzyme concentrations of 12.5, 50 and 80 $\mu\text{g/ml}$ respectively. In each case the limiting regain of activity (after 24 h) has been normalized to 100, and the continuous lines represent the fit to a second-order model with a rate constant of 900 $\text{M}^{-1} \cdot \text{s}^{-1}$. The inset shows the actual extent of regain of activity (ordinate) as a function of the concentration of enzyme (abscissa) relative to a non-denatured control at that enzyme concentration.

ml. The kinetics of the regain of activity are shown in Figure 2, where, for the ease of comparison, the data have been normalized to the limiting extent of re-activation (observed after 24 h) as 100 % in each case. In each case it is clear that there is a short lag, followed by one or more processes which show second-order characteristics, in that the half-time is inversely proportional to the concentration of enzyme. The curves shown in Figure 2, which give a reasonably satisfactory fit to the observed data, were calculated assuming a second-order rate constant of 900 $\text{M}^{-1} \cdot \text{s}^{-1}$. The possible nature of this associative process is considered in the Discussion section.

The regain of activity of the *S. coelicolor* enzyme was found to depend on the presence of dithiothreitol during both the unfolding and refolding steps. Omission of dithiothreitol from both steps resulted in the regain of < 5 % of the activity observed if the reducing agent was included in both. Omission of dithiothreitol from the refolding step only resulted in regain of about half of the activity observed if it were included in both steps. This limited regain of activity was presumably a result of the carry-over of the reducing agent from the unfolding step. The requirement of dithiothreitol for the regain of activity reflects the fact that the *S. coelicolor* enzyme has four cysteine side chains per subunit, which would be likely to be subject to oxidative damage during the unfolding and refolding procedures.

In the case of the *M. tuberculosis* enzyme, the regain of activity was much less efficient than in the case of the *S. coelicolor* enzyme, even when dithiothreitol was included in the unfolding and refolding steps. At protein concentrations (after dilution) of 10, 22 and 80 $\mu\text{g/ml}$, the average limiting extents of reactivation after 120 h were 13, 30 and 2 % respectively, indicating that aggregate formation may be a competing process, especially at the higher concentrations. In addition, it should be noted that there was considerable variation (up to 20 %) between replicate experiments, making it difficult to analyse the kinetics in detail. However, it was clear in all cases that the rate of recovery of activity showed a marked dependence on the concentration of enzyme (with the half times for re-activation being approx. 100

Figure 3 Refolding of the type II DHQases from *S. coelicolor* at 80 $\mu\text{g/ml}$ monitored by stopped-flow CD measurements

The traces A, B and C refer to enzyme in 6 M GdmCl, enzyme in 0.55 M GdmCl and refolding enzyme respectively. Each trace is the average for four sets of data, each consisting of five shots. The continuous line shown in C represents a first-order fit, with a rate constant of 0.07 s^{-1} . Abbreviation: mdeg, millidegrees.

and 400 min at concentrations of 22 $\mu\text{g/ml}$ and 10 $\mu\text{g/ml}$ respectively), consistent with a process limited by one or more associative steps. In view of the difficulties experienced with the *M. tuberculosis* enzyme, all subsequent detailed measurements were performed on the *S. coelicolor* enzyme.

Changes in CD during refolding

Using manual mixing techniques, the regain of the secondary structure (ellipticity at 230 nm) on refolding at a concentration of 80 $\mu\text{g/ml}$ was found to be more than 80 % complete within 60 s of the start of refolding. The far-UV CD spectrum of the enzyme after 15 min refolding was very similar in shape to that of non-denatured enzyme over the range 260–220 nm, with approx. 85 % of the amplitude, indicating that the main secondary structural features of the enzyme had been regained. It was not possible to record data below 220 nm because of the presence of the residual GdmCl (0.55 M), which leads to high absorbance in the 0.5-cm-pathlength cuvette, which was necessary at this protein concentration.

The stopped-flow CD experiments at 80 $\mu\text{g/ml}$ enzyme showed that 90 % of the native ellipticity at 230 nm was regained on refolding. About $40 \pm 5\%$ of this change was complete within the dead time of the instrument in the CD mode (8 ms); the remaining change took place in a first-order fashion with a rate constant of $0.07 \pm 0.01\text{ s}^{-1}$ (Figure 3). The CD experiment was repeated at a concentration of 43 $\mu\text{g/ml}$ in the refolding mixture, and essentially identical results were obtained (85 % regain of native ellipticity; 50 % of this within the dead time, the remainder with a rate constant of 0.06 s^{-1}). It was not possible to observe the refolding by stopped-flow CD at lower concentrations of enzyme because of the high level of noise associated with such measurements.

Changes in fluorescence during refolding

Upon refolding of the *S. coelicolor* enzyme, the fluorescence emission maximum (excitation at 290 nm) shifts back to 341 nm, a value identical with that of the enzyme in the presence of 0.55 M GdmCl and very close to that of the native enzyme

Figure 4 Kinetics of changes in fluorescence of the type II DHQase from *S. coelicolor* during refolding

Curves A, B and C represent enzyme concentrations of 12.5, 50 and 80 $\mu\text{g/ml}$ respectively. Each trace has been normalized to a limiting value of 100 as that observed after 30 min refolding. The continuous lines represent the fit to a second-order model with a rate constant of $5700\text{ M}^{-1}\text{ s}^{-1}$. The inset shows the actual regain of fluorescence as a function of the concentration of enzyme relative to a non-denatured control at that enzyme concentration,

(340 nm). The time course of changes in fluorescence at 350 nm during refolding are shown in Figure 4, normalized to the total change observed over a 30 min period as 100 %. There was no further change over a further 15 min, though on long-term observation a slow decline in fluorescence (approx. 10 %/h) was observed, presumably reflecting photochemical damage to the sample. The amplitude of the fluorescence change observed follows a similar trend to that shown by the regain of activity (see Figure 2) declining from 77 % at a protein concentration of 80 $\mu\text{g/ml}$ to 43 % at 12.5 $\mu\text{g/ml}$. Inspection of the kinetic data shows that, at 80 $\mu\text{g/ml}$, the change during the dead time of the manual mixing experiment (10 s) is about 25 % of the total; this decreases to less than 10 % at 12.5 $\mu\text{g/ml}$. The remaining fluorescence change occurs in a second-order fashion, with a satisfactory fit to the data shown by assuming a rate constant of $5700\text{ M}^{-1}\text{ s}^{-1}$, a value which is comparable with the rate constants for subunit association processes measured during the refolding of a number of oligomeric enzymes [1]. The nature of this associative process is considered further in the Discussion section. It is important to note that the fits to the regain of activity (Figure 2) and of fluorescence (Figure 4) are both made using simple second-order models. It is probable that more complex models involving sequential second-order or second- and first-order processes would give more precise fits to the experimental data, but in our view such complex kinetic modelling is not justified in view of the likelihood that a number of complex models could give equally good fits to the observed data.

When the fluorescence changes during refolding were examined using stopped-flow mixing, the size of the fluorescence change within the first 10 ms of the refolding reaction was found to be less than 15 % of the total change, even at the highest concentration studied (80 $\mu\text{g/ml}$), with the remaining part of the change fitted in a very satisfactory manner as a second-order process.

The fluorescence properties of the species formed in the early stages of the refolding process were examined by manual mixing followed by rapid scanning (100 nm/min) of the fluorescence over the wavelength range from 330 to 370 nm. Under these

Figure 5 Comparison of the rates of regain of CD, fluorescence and activity during the refolding of type II DHQase from *S. coelicolor* at an enzyme concentration of 80 $\mu\text{g/ml}$

For each parameter the value is scaled to show the percentage of the total limiting change observed under those conditions. The shift in fluorescence emission maximum during refolding is also indicated.

conditions, it was found that the wavelength of the emission maximum could be reliably determined within 35 ± 5 s of the start of the refolding process. At a concentration of 80 $\mu\text{g/ml}$ during refolding, the emission maximum after 35 s was 349 nm; over the course of the subsequent 30 min, the maximum moved towards 341 nm (Figure 5). The experiment was repeated at a lower concentration of protein during refolding (12.5 $\mu\text{g/ml}$) so that the rate of second-order associative processes would be slowed sufficiently to allow a more accurate measure of the fluorescence properties of initially formed folded monomer to be made. Under these conditions, the wavelength maximum 35 s after refolding was initiated was 351 nm; over the subsequent 30 min, the maximum moved to 345 nm. Attempts to use the stopped-flow instrument to measure the shift in the emission maximum in the early phases of the refolding process were not successful because, under the conditions employed, the signal-to-noise ratio was not sufficient to allow accurate estimation of the emission maximum.

Reassembly of the enzyme on refolding

The changes in apparent molecular mass on refolding were monitored by gel-permeation chromatography. When enzyme which had been incubated for 6 h in 6 M GdmCl was diluted with 10 vol. of buffer to initiate refolding (to give final concentrations of enzyme and GdmCl of 80 $\mu\text{g/ml}$ and 0.55 M respectively), the apparent molecular mass observed after 1 h (50 kDa) corresponded to that of the trimer. A small peak corresponding to about 10% of the total material was eluted at a volume corresponding to that of the subunit (16 kDa); this material was presumed to represent residual monomeric forms of the enzyme which had not folded to a form that could assemble into oligomers. In view of the finding that the dodecameric enzyme was dissociated into trimers at concentrations of GdmCl even as low as 0.2 M, it was clearly not practicable to study the reassembly of the intact dodecamer by using dilution to initiate refolding. Instead, extensive dialysis of the enzyme incubated with GdmCl was used to lower the concentration of the dena-

turing agent to below 0.01 M. When this procedure was applied to 80 $\mu\text{g/ml}$ enzyme which had been incubated in 6 M GdmCl, over 90% of the enzyme was found to be eluted at the volume corresponding to the dodecamer, with the remainder corresponding to the subunit molecular mass; under these conditions 95% of the activity of native enzyme was regained. A similar procedure applied to enzyme which had been incubated in 0.55 M GdmCl and then dialysed yielded exclusively the dodecameric enzyme. These experiments confirmed that it was possible to reassemble the intact active, dodecameric, *S. coelicolor* enzyme in high yield after denaturation in GdmCl.

DISCUSSION

From the data obtained on the unfolding of the two DHQases by GdmCl, it is clear that dissociation of the dodecameric enzyme into trimeric units can be brought about at very low concentrations (≤ 0.2 M) of GdmCl. This dissociation involves no significant change in the far-UV CD or fluorescence properties of the enzyme, indicating that the secondary and tertiary structures of the subunits are essentially unaltered. As judged by the data obtained at 0.5 M GdmCl, the trimeric units possess 85 or 155% of the activity of the native enzymes from *S. coelicolor* or *M. tuberculosis* respectively (Figure 1). The dissociation behaviour of the enzymes can be understood in terms of the X-ray structures of the *M. tuberculosis* and *S. coelicolor* enzymes. In both cases, the dodecameric enzyme consists of four trimers arranged with cubic 23 symmetry. Within the trimeric units there is an extensive network of mainly side-chain interactions, which are different in number and type between the two enzymes; these interactions bury $\approx 12 \text{ nm}^2$ ($1200 \text{ \AA}^2$) of surface-accessible area per subunit. At each interface between the subunits in the trimer there are three salt bridges and nine hydrogen bonds in the case of the *M. tuberculosis* enzyme and one salt bridge and seven hydrogen bonds in the case of the *S. coelicolor* enzyme. For both enzymes the interactions between the trimeric units are less extensive than those within the trimers. On formation of the dodecameric structures of the *M. tuberculosis* and *S. coelicolor* enzymes 9.5 and 8 nm^2 (950 and $800 \text{ \AA}^2$) of surface-accessible area are buried, with the formation of ten and nine hydrogen bonds respectively. No salt bridges are involved in either enzyme. The pattern of more extensive interactions within the trimer than between trimers is consistent with the proposal made in the present paper that the trimer represents the minimum active unit of the type II DHQases. It should be noted that, by contrast, there are extensive interactions between the trimeric units of ornithine carbamoyl-transferase from *P. aeruginosa* and *P. furiosus*, contributing to the stability of the dodecameric structures of these enzymes [13,15].

The location of the active site of each subunit of the type II DHQases can be deduced from the pattern of conserved residues and from the results of chemical modification and site-directed-mutagenesis experiments, which indicate the importance of Arg²³ and Tyr²⁸ [23,24]. From the X-ray structure it appears that the catalytically important amino acids at each active site are contributed by that subunit alone. The necessity for the trimer as the minimum catalytically active unit may well be explained by the fact that the side chain of Arg¹¹⁷ (*S. coelicolor*) or Arg¹¹² (*M. tuberculosis*) is involved in an intersubunit salt bridge; this maintains the orientation of one of the loops which contribute to the architecture of the active site (Figure 6).

From the kinetic data reported here, it is clear that the refolding of *S. coelicolor* DHQase occurs in a number of discrete phases; the relative rates of these are compared in Figure 5, and the processes concerned are summarized in the outline scheme

Figure 6 Structure of the type II DHQase from *S. coelicolor*

The Figure shows a subunit of the enzyme (grey) and parts of the two other subunits (green and blue) within the trimer; a sulphate ion is bound in the centre of the three-fold symmetry. Active-site residues Arg²³, Tyr²⁸, His⁸⁵, His¹⁰⁶ and Arg¹¹⁷ are shown in yellow. The active-site residue Arg¹¹⁷ forms a salt bridge with Asp⁹² from a subunit related by three-fold symmetry. The position of Trp⁶⁵ in each subunit is also indicated. The Figure was generated using the program MOLSCRIPT [26].

shown in Table 1. In the first phase about half of the secondary structure is regained very rapidly (within 10 ms) as judged by the size of the CD signal at 230 nm; the remaining secondary structure is regained in a rather slower first-order process. These first-order processes, which presumably represent changes occurring within a single subunit, are accompanied by only a small fraction of the total change in the fluorescence signal at 350 nm. This indicates that, in this folded monomeric intermediate, the environment of the single tryptophan side chain is not greatly different from that of the unfolded enzyme, a conclusion supported by the finding that the emission maximum (351 nm) observed in the early phases of the refolding of the protein at 12.5 µg/ml is similar to that of the unfolded enzyme (356 nm). Following the completion of the first-order processes, most of the increase in the fluorescence signal occurs with second-order kinetics (Figure 4), showing that at least one associative event is

required to bring about this change. This association of subunits is accompanied by a blue shift in the emission maximum (to 341 nm), consistent with the partial burial of the tryptophan side chain in each subunit (Figure 5). It should be noted that, at the earliest time point (35 s) at which the emission maximum could be reliably estimated, some association of subunits would have already occurred, so it is probable that the emission maximum of a folded monomeric intermediate is actually slightly higher than 351 nm.

Regain of activity occurs even more slowly, with a small lag period (Figures 2 and 5); thus, for instance, at a concentration of 25 µg/ml after 10 min refolding, over 90% of the limiting fluorescence change had occurred, whereas the re-activation was only 44% of the limiting value. It is therefore likely that the second-order process observed (after the lag phase) for regain of activity reflects a further associative step to generate an active

Table 1 Refolding of type II DHQase from *S. coelicolor*

In the Table, U and N refer to the unfolded and native forms of the enzyme respectively. M_1 , M_2 , D and T refer to two forms of the monomer, the dimer and trimer respectively. The first-order rate constant for conversion of M_1 to M_2 is 0.07 s^{-1} , and the second-order rate constants for conversion of M_2 to D and for D to T are 5700 and $900\text{ M}^{-1}\text{ s}^{-1}$ respectively. In this Table the values for CD at 230 nm, fluorescence at 350 nm and activity represent the percentage of the value of the native form (N), with the values for U being set at zero. For further details, see the text.

	U	M_1	M_2	D	T	N
CD (230 nm)	0	40	100	100	100	100
Fluorescence (350 nm)	0	10	10	100	100	100
Fluorescence λ_{max} (nm)	356	≥ 351	≥ 351	341	341	340
Activity	0	0	0	0	85	100

species. Thus, from the finding that two consecutive second-order processes are required for the regain of activity, it can be proposed that the minimum size of the active unit is a trimer. This suggestion would be consistent not only with the X-ray structures of the two enzymes, but also with the gel-permeation data, which show that both enzymes are dissociated into trimers at low concentrations of GdmCl.

The X-ray structures of the enzymes (D. G. Gourley, A. K. Shrive, I. Polikarpov, T. Krell, A. R. Hawkins, L. Sawyer, N. W. Isaacs and J. R. Coggins, unpublished work; A. Roszak, D. G. Gourley, T. Krell, J. R. Coggins and A. J. Laphorn, unpublished work) show that the single tryptophan residue in each subunit is relatively close to the subunit interface within the trimer (Figure 6). Thus, in the putative folded monomeric intermediate, it would appear that the tryptophan side chain would be reasonably exposed to solvent, consistent with the fluorescence emission maximum above 350 nm. The subsequent change in the observed fluorescence occurs in a second-order process (see Figure 4), reflecting the change in the tryptophan environment on association of the subunits.

The trend in yield of regain of activity with concentration of the *S. coelicolor* enzyme is in line with data from other oligomeric proteins [1], namely at low concentrations of protein, the isolated 'folded' subunits are unstable and effectively removed from the system before productive association and stabilization reactions can occur. At all concentrations studied the yield is less than 100%, indicating the occurrence of side reactions, the most likely of which are misfolding of individual subunits. Formation of aggregates is possible, but these have not been observed to any significant extent in our gel-permeation experiments. By contrast, in the case of the *M. tuberculosis* enzyme, the low yield of activity at a concentration of $80\text{ }\mu\text{g/ml}$ suggests that, under these conditions, aggregate formation is dominant.

We have employed several strategies to measure the time course of regain of quaternary structure of the *S. coelicolor* enzyme during refolding and thus to be able to confirm the nature of the intermediates suggested by the kinetic analysis (Table 1). However, none of these has proved successful. The cross-linking approach using glutaraldehyde [25] does not yield any high-molecular-mass product; this lack of reaction is presumably due to the fact that there is only one lysine residue per subunit. Measurements of light-scattering changes during refolding at a protein concentration of $80\text{ }\mu\text{g/ml}$ have proved inconclusive, even when considerable care has been taken to remove dust and other particles greater than $0.2\text{ }\mu\text{m}$ by Millipore filtration. The differences in amplitudes between native, unfolded and refolding enzyme have been small and irreproducible. At the

concentrations used in these experiments, no other techniques currently available would be able to monitor changes in molecular mass over the required time scale. Nevertheless, the finding that, in the presence of 0.55 M GdmCl, the enzyme is trimeric, has allowed attention to be focused on the steps involved in trimer formation as the GdmCl concentration is lowered from 6 M to 0.55 M by dilution. Our studies have allowed an outline pathway for the folding and assembly of the enzyme to be proposed, which can serve as a basis for more detailed investigations using a combination of protein-engineering and biophysical methods. Using a dilution approach it is not practicable to study the complete folding and assembly of the subunits into dodecamers, since the residual concentration of GdmCl which would permit dodecamer formation would be very low. However, by using dialysis to remove the denaturing agent, full reassembly can be demonstrated.

The comparative structural data for the two enzymes (D. G. Gourley, A. K. Shrive, I. Polikarpov, T. Krell, A. R. Hawkins, L. Sawyer, N. W. Isaacs and J. R. Coggins, unpublished work; A. Roszak, D. G. Gourley, T. Krell, J. R. Coggins, and A. J. Laphorn, unpublished work) may help to shed some light on the reasons for the lower yields of re-activation for the enzyme from *M. tuberculosis* compared with that from *S. coelicolor*. The X-ray data show that the interactions between the subunits of the trimer are more extensive in the case of the former enzyme. Therefore association of subunits to give the active trimeric unit of the *M. tuberculosis* enzyme would require a greater number of precise interactions and thus be likely to occur with a decreased yield. In the absence of additional folding factors the initial folded subunits would have a tendency to aggregate if the native intersubunit interactions did not occur, especially at high protein concentrations. The folding and assembly of these complex dodecameric enzymes *in vivo* would presumably involve the intervention of chaperone proteins to improve both the rate and amplitude of the successful regain of activity.

We thank the Biotechnology and Biological Sciences Research Council for financial support, and Dr. Adrian Laphorn for valuable discussions.

REFERENCES

- Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol.* **49**, 117–237
- Gething, M.-J. and Sambrook, J. (1992) *Nature (London)* **355**, 33–44
- Pain, R. H. (1994) *Mechanisms of Protein Folding*, Oxford University Press, Oxford
- Hartl, F. U. (1996) *Nature (London)* **381**, 571–580
- Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S. and Nagamura, T. (1987) *FEBS Lett.* **221**, 115–118
- Goldberg, M. E., Semisotnov, G. V., Friguet, B., Kuwajima, K., Ptitsyn, O. B. and Sugai, S. (1990) *FEBS Lett.* **263**, 51–56
- Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M. and Poulsen, F. M. (1995) *Biochemistry* **34**, 7217–7224
- Dobson, C. M., Evans, P. A. and Radford, S. M. (1994) *Trends Biochem. Sci.* **19**, 31–37
- Mann, C. J. and Matthews, C. R. (1993) *Biochemistry* **32**, 5282–5290
- Price, N. C. (1994) in *Mechanisms of Protein Folding* (Pain, R. H., ed.), pp. 160–193, Oxford University Press, Oxford
- Hlodan, R. and Hartl, F. U. (1994) in *Mechanisms of Protein Folding* (Pain, R. H., ed.), pp. 194–228, Oxford University Press, Oxford
- Kleanthous, C., Dekka, R., Davis, K., Kelly, S. M., Cooper, A., Harding, S. E., Price, N. C., Hawkins, A. R. and Coggins, J. R. (1992) *Biochem. J.* **282**, 687–695
- Villeret, V., Tricot, C., Stalon, V. and Dideberg, O. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10762–10766
- Nguyen, V. T., Baker, D. P., Tricot, C., Baur, H., Villeret, V., Dideberg, O., Gigot, D., Stalon, V. and Haas, D. (1996) *Eur. J. Biochem.* **236**, 283–293
- Villeret, V., Clantin, B., Tricot, C., Legrain, C., Roovers, M., Stalon, V., Glansdorff, N. and van Beeumen, J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2801–2806
- Burnett, R. M., Darling, G. D., Kenall, D. S., Le Quesne, M. E., Mayhew, S. G., Smith, W. W. and Ludwig, M. (1974) *J. Biol. Chem.* **249**, 4383–4392
- White, P. J., Young, J., Hunter, I. S., Nimmo, H. G. and Coggins, J. R. (1990) *Biochem. J.* **265**, 735–738

- 18 Moore, J. D., Lamb, H. K., Garbe, T., Servos, S., Dougan, G., Charles, I. G. and Hawkins, A. R. (1992) *Biochem. J.* **287**, 173–181
- 19 Moore, J. D., Hawkins, A. R., Charles, I. G., Deka, R., Coggins, J. R., Kelly, S. M. and Price, N. C. (1993) *Biochem. J.* **295**, 277–285
- 20 Nozaki, Y. (1972) *Methods Enzymol.* **26**, 43–50
- 21 Chaffotte, A. F., Cadieux, C., Guillou, Y. and Goldberg, M. E. (1992) *Biochemistry* **32**, 4303–4308
- 22 Tonomura, B., Nakatani, H., Masatake, O., Junko, Y.-I. and Horomi, K. (1977) *Anal. Biochem.* **84**, 370–383
- 23 Krell, T., Pitt, A. R. and Coggins, J. R. (1995) *FEBS Lett.* **360**, 93–96
- 24 Krell, T., Horsburgh, M. J., Cooper, A., Kelly, S. M. and Coggins, J. R. (1996) *J. Biol. Chem.* **271**, 24492–24497
- 25 Jaenicke, R. and Rudolph, R. (1986) *Methods Enzymol.* **131**, 218–250
- 26 Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950

Received 4 September 1998/3 November 1998; accepted 23 November 1998