Structural Characterization of the M* Partly Folded Intermediate of Wild Type and P138A Aspartate Aminotransferase from Escherichia coli

Leila Birote, Fabrizio Dal Piaz, Piero Pucci, and Gennaro Marino

From the Dipartimento di Chimica Organica e Biochimica, Universita` Federico II di Napoli, Complesso Universitario Monte Sant’Angelo, Via Cintia, 80126 Napoli, Italy

A combination of spectroscopic techniques, hydrogen/deuterium exchange, and limited proteolysis experiments coupled to mass spectrometry analysis was used to depict the topology of the monomeric M* partly folded intermediate of aspartate aminotransferase from Escherichia coli in wild type (WT) as well as in a mutant form in which the highly conserved cis-proline at position 138 was replaced by a trans-alanine (P138A). Fluorescence analysis indicates that, although M* is an off-pathway intermediate in the folding of WT aspartate aminotransferase from E. coli, it seems to coincide with an on-pathway folding intermediate for the P138A mutant. Spectroscopic data, hydrogen/deuterium exchange, and limited proteolysis experiments demonstrated the occurrence of conformational differences between the two M* intermediates, with P138A-M* being conceivably more compact than WT-M*. Limited proteolysis data suggested that these conformational differences might be related to a different relative orientation of the small and large domains of the protein induced by the presence of the cis-proline residue at position 138. These differences between the two M* species indicated that in WT-M* Pro138 is in the cis conformation at this stage of the folding process. Moreover, hydrogen/deuterium exchange results showed the occurrence of few differences in the native N2 forms of WT and P138A, the spectroscopic features and crystallographic structures of which are almost superimposable.

The mechanism by which proteins fold into their unique native structures is still a central problem in structural biology. It is increasingly recognized that the structure of non-native states of proteins can provide significant insight into fundamental issues such as the relationship between protein sequences and three-dimensional structures, the nature of protein folding pathways, the stability of proteins and their turnover in the cell, and the transport of proteins across membranes (1). Moreover, intermediate states experienced by proteins in vivo often play a major role in protein association and aggregation, leading to the “so-called” conformational diseases (2). In contrast to the large amount of structural information available on native folded proteins, however, too few partly folded intermediates have been characterized thoroughly enough to propose general models on how the native state is attained and which is the structure of transient folding states. Investigation of a wide range of non-native states would then be of considerable value. To meet this need, new analytical strategies able to characterize transient species and to define the molecular details through which diverse proteins fold are required.

Recently, structural biologists have turned their attention to integrated strategies for the definition of the surface topology of proteins and protein complexes. Although these approaches provide low resolution data, they are amenable to the analysis of transient species and partly folded intermediates. Limited proteolysis and amide hydrogen exchange experiments in combination with mass spectrometry have been employed to investigate the surface topology and the conformational flexibility of proteins and protein complexes. Amide protons are exchanged with deuterium with kinetics mainly depending on both solvent accessibility and stabilization of the associated protein backbone region (3–5). Comparative hydrogen/deuterium exchange (H/D)1 measurements can then be used to monitor protein structural changes in different experimental conditions (6), as well as the effects of binding or aggregation (7). Although hydrogen exchange has most often been measured by NMR, monitoring by mass spectrometry (MS) has become increasingly common (8–13). As deuterium atoms replace protons during the hydrogen exchange reaction, the mass of the protein increases. The extent as well as the rate of exchange can then be determined by measuring the increase in protein mass.

Proteolytic cleavages on a protein substrate can occur only if the polypeptide chain is exposed and flexible enough to adapt to the specific protease active site. The stable conformation of proteins then provides some stereochemical barriers to enzymatic attack, leaving the exposed and flexible regions accessible to proteases and preventing the occurrence of proteolytic cleavages within the highly structured core of the molecule. Consequently, when these experiments are performed using a series of proteases with different specificity under conditions

1 The abbreviations used are: H/D, hydrogen/deuterium exchange; ExAspAT, aspartate aminotransferase from E. coli; WT, wild type aspartate aminotransferase from E. coli; P138A, aspartate aminotransferase from E. coli mutant form in which the cis-proline at position 138 was replaced by a trans-alanine; mAspAT, rat liver mitochondrial aspartate aminotransferase; CD, circular dichroism; WT-N2, native form of wild type; P138A-N2, native form of P138A; WT-M*, M* intermediate from wild type; P138A-M*, M* intermediate from P138A; ANS, 8-anilino-1-naphthalene-1-sulfonic acid; GdmHCl, guanidinium hydrochloride; PLP, pyridoxal 5′-phosphate; PMP, pyridoxamine 5′-phosphate; HPLC, high performance liquid chromatography; DTT, dithiothreitol; ES, electrospray; MS, mass spectroscopy; LC, liquid chromatography.
that favors a single bond cleavage (complementary proteolysis),
the pattern of preferred cleavage sites will depict the exposed
regions in the protein molecule (14, 15). This strategy was also
employed to investigate conformational changes occurring in
protein structure under different experimental conditions (16–
19) or during quaternary forms interchange (20, 21), as well as
for the definition of the interface regions in protein complexes
(22–24).

This paper reports the application of these integrated strat-
egies to the structural characterization of partly folded inter-
mediates of wild type and a mutant form of Escherichia coli
aspartate aminotransferase (EcAspAT), in which the highly
conserved cis-proline in position 138 had been replaced by a
trans-alanine (P138A; Ref. 25). EcAspAT is a homodimer with
two identical and independent active sites located at the sub-
unit interface. Each subunit consists of an N-terminal arm, a
large cofactor-binding domain, and a small domain (26). Pro-
138 is located in a loop region of the large domain close to the
interface with the small domain (27). The replacement of a
cis-proline by a trans-alanine did not significantly affect either
the activity or the stability of the protein, as the catalytic
efficacy of P138A is of the same order of magnitude of WT,
and the thermal unfolding curves of the mutant and the WT are
almost superimposable (25). Additionally, the GdmHCl-in-
duced unfolding equilibrium of P138A mutant follows the same
pathway as the wild type protein (25). Nevertheless, the mu-
tant enzyme shows interesting folding features, because, de-
spite the replacement of a cis peptide bond with a trans one, the
refolding process is slower than wild type (25). Moreover, the
monomeric P138A-M* intermediate (25) detectable in the
GdmHCl-induced unfolding at equilibrium exhibited different
spectroscopic properties as compared with the wild type M* species and resembles a kinetic folding intermediate.

**EXPERIMENTAL PROCEDURES**

**Materials**—Recombinant wild type and P138A mutant EcAspATs were isolated from the overproducing E. coli strain TY103 (28) as
described previously (25). The proteins were stored at —80 °C in
the presence of 1 mM DTT and an excess of PMP and 2-oxoglutarate.

All experiments were carried out on the PPL form of the enzymes,
obtained by removal of excess PMP and 2-oxoglutarate either by exten-
sive dialysis or by appropriate high-speed gel filtration on a Super-
ose 12 PC column (3.2 × 30 mm) using the SMART system (Amersham
Biosciences). Subunit concentration of EcAspATs was determined spec-
trophotometrically on a Beckmann DU7500 spectrophotometer, using
ε280 = 49,935 M⁻¹ cm⁻¹ (29).

1-1-Tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin,
chymotrypsin, and subtilisin were purchased from Sigma; endoprotease
Lys C was from Roche Molecular Biochemicals. Reverse phase-HPLC
C4 columns (25 × 0.46 cm, 5 mm) were purchased from Vydac (The
Separation Groups). All other reagents and solvents were HPLC-grade
from Carlo Erba.

**Spectroscopic Analysis**—Native EcAspAT holoenzymes (in 10 mM
HEPES, pH 7.5, 1 mM DTT, and 0.15 M NaCl) were incubated at 25 °C
at the required protein and GdmHCl concentrations for 1 h (folding/
unfolding equilibrium was attained within 30 min, as determined in
preliminary experiments by following by time the enzymatic activity
and fluorescence), and protein solutions were filtered just before the
analysis on 0.22-μm pore size polyvinylidene difluoride membrane
(Millipore).

Far-UV CD spectra were recorded on a Jasco J715 spectropolarime-
ter equipped with a Peltier thermostatic cell holder (Jasco model PTC-
348), in a quartz cell (0.1-cm light path) at a protein concentration of 1.0
μM. Temperature was measured directly in the quartz cell, the solutions
were filtered just before use on 0.22-μm pore size polyvinylidene diflu-
oride membrane (Millipore), and data corrected by subtracting a control
from which the protein was omitted. Spectra were recorded at 25 °C
from 280 to 200 nm at 0.2-nm resolution, 16-s response, at a scan rate
of 20 nm/min. All data are the averages of three measures, and the
results are expressed as mean residue ellipticity [θ], which is defined as

\[
\theta = -100 \Delta \alpha_{222}/c, \quad \text{where } \Delta \alpha_{222} \text{ is the observed ellipticity in degrees, } c \text{ is the}
\]
concentration in molar units, and l is the length of the light path in
centimeters.

Fluorescence measurements were carried out on a PerkinElmer
LB508 fluorimeter, using an optical cuvette of 10-mm light path length
with thermostatically controlled cell holder. Tryptophan emission spec-
tra were measured directly in the quartz cell, at a protein concentration of 1.0 μM using an excita-
tion wavelength of 295 nm, with excitation and emission bandwidths
of 10 and 2.5 nm, respectively. Tryptophan emission spectra were re-
corded between 100 and 200 nm at a scan rate of 100 nm/min. Each
spectrum is the average of three emission scans, and data were cor-
rected by subtracting a blank from which the enzyme was omitted. ANS
fluorescence spectra were recorded using an excita-
tion wavelength of 375 nm with excitation and emission bandwidths
of 5 and 5 nm, respectively between 400 and 600 nm at a scan rate of
100 nm/min. Each spectrum is the average of three emission scans, and
data were corrected by blank minus spectrum.

Single-jump refolding experiments were carried out by concentration
jumps of GdmHCl at 5 °C by rapid dilution of protein previously incu-
bated at required higher denaturant concentrations to the desired one
at a final protein concentrations of 0.39 μM, and the subsequent time-
dependent changes in tryptophan or ANS fluorescence intensity were
monitored, using a LB50S fluorimeter, using an optical cuvette of 10-mm light path length.

**Tryptophan Fluorescence Quenching**—Fluorescence quenching titra-
tions with either acrylamide or iodoide were performed at 25 °C by
sequential addition of aliquots of concentrated acrylamide (8 μM) or
potassium iodoide (5.34 μM) stock solution into the protein solution.
For the native state, protein concentration was 3.9 μM in 10 mM Na-HEPES,
ph 7.5, 1 mM DTT, and 0.15 M NaCl in a range of acrylamide concentration
from 0 to 1.03 M. For the M* intermediate, iodoide was added to the protein
(0.39 μM previously incubated at 25 °C for 30 min in 10 mM
HEPES, pH 7.5, 1 mM DTT, and 0.15 M NaCl at 1.3 M GdmHCl for
P138A-M*, and 1.5 M GdmHCl for WT-M*) to obtain final iodoide con-
centration ranging from 0 to 0.2 μL. Fluorescence was measured at 25 °C
at the wavelength of the maximum in the fluorescence spectrum re-
corded in the absence of the quenching agent. Excitation was at 295 nm
(10-nm excitation bandwidth, 2.5-nm emission bandwidth), corrected for
the minor dilution caused by addition of the quenching agent and for
the background fluorescence of both solvent and quenching agent.
Correction for the inner filter effect that resulted from the absorption of
acrylamide at 295 nm was applied by multiplying the fluorescence
intensity by 10⁻⁴, where A is the absorbance of the solution at 295 nm
measured in a 1-cm path length cuvette. In the iodoide experiments, 0.1
M Na₂S₂O₃ was included to prevent I⁻ ion formation.

The fluorescence-quenching data in the presence of iodoide quenching
data were analyzed according to the Stern-Vollmer equation (Equation (1))
(30).

\[
F/F_0 = 1 - K_{q} [Q]
\]

(Eq. 1)

F₉ and F₀ are the fluorescence intensities in the absence and presence
of quencher, respectively, Kₚₗ is the Stern-Vollmer quenching constant,
and [Q] is the concentration of the quencher. The plot of F/F₀ versus [Q]
is linear for an apparently homogeneous population of emitting fluoro-
phores. Alternatively, a Stern-Vollmer-derived equation (Equation (2))
was used to calculate the fraction of accessible fluorescence, that is, the
fraction of tryptophans that are exposed and then accessible to the
quenching agent.

\[
F/F_0 = 1 - K_{q} [Q] \exp \left( V/\lambda \right)
\]

(Eq. 3)

V is the static constant.

Gel Filtration—Gel filtration experiments were carried out as fol-
ows. 50 μL of the samples, 10 μM in 10 mM HEPES, pH 7.5, 1 mM
DTT, and 0.15 M NaCl, equilibrated for 30 min at different concentrations
of GdmHCl, were loaded on a Superose 12 PC (3.2 × 300 mm) gel filtration
column installed on the Smart System (Amersham Biosciences), and
isocratically eluted in 10 mM HEPES, pH 7.5, 1 mM DTT, and 0.15 μL
NaCl at the same concentration of GdmHCl used in the equilibration.
All experiments were carried out at 25 °C at a flow rate of 40 μl/min.
A Partly Folded Intermediate of EcAspAT

The column was calibrated in 10 mM HEPES, pH 7.5, 1 mM DTT, and 0.15 mM NaCl in the absence of GdmHCl with the following proteins of known molecular mass: β-amylose (200,000 Da), cAspAtp (92,000 Da), bovine hemoglobin (67,000 Da), cytochrome e (12,300 Da).

Hydrogen/Deuterium Exchange—The H/D reaction was conducted as follows: proteins were dissolved in 50 mM HEPES, pH 7.5, 1 mM DTT, and 0.15 mM NaCl at 25 °C, in the presence of 1.4 M GdmHCl in the case of the two M* species, at a 8 μM protein concentration, and allowed to equilibrate for 30 min. Deuterium exchange reactions were initiated by 10-fold dilution of the sample with 50 mM HEPES, pH 7.5, D_2O (containing 1.4 M GdmHCl in the case of the two M* species). The exchange reaction was allowed to proceed for lengths of time ranging from 30 to 60 min, and, at each time point, 1 nmol of protein was removed from the labeling solution and rapidly injected into a 30 mm × 0.46 mm (inner diameter) perfusion column (POROS R2 media, Applied Biosystems) on an HPLC equipped with two Series 200 LC isocratic pumps (Applied Biosystems) coupled to an API-100 single quadrupole mass spectrometer (Applied Biosystems). The protein was eluted at a flow rate of 1 mL/min with a gradient of 25–95% acetonitrile in 0.1% trifluoroacetic acid in 1.0 min. The HPLC step was performed with cold protiated solvents, thereby reducing the back-exchange kinetics and removing deuterium from side chains and N/C termini that exchange much faster than amide linkages (32). Therefore, the increase in molecular mass provided a direct measurement of the deuteration at peptide amide linkages. Data were acquired and elaborated using the Biomultiviewer (Applied Biosystems) program. Non-deuterated and totally deuterated samples were used as control. The fully deuterated samples were prepared by incubation for 1 h at 25 °C in 50 mM HEPES, pH 7.5, 1 mM DTT, 0.15 mM NaCl, 5 mM GdmHCl in D_2O at 8 μM protein concentration. LC/MS analysis was performed as above.

Limited Proteolysis Experiments—Limited proteolysis experiments were carried out by incubating P138A-M*, WT-M*, P138A-N2, and WT-N2 with trypsin, chymotrypsin, subtilisin, and endoprotease Lys C using enzyme to substrate ratios ranging from 1:1000 to 1:5 (w/w). Enzymatic digestions were all performed in 50 mM HEPES, pH 7.5, 1 mM DTT, and 0.15 mM NaCl at 25 °C, in the presence of the desired concentration of GdmHCl (1.3 M for P138A-M*, 1.5 M for WT-M*), and without GdmHCl for the two native forms, at a 1 μM protein concentration. In the case of P138A-M* and WT-M*, each protein was incubated for 30 min in the reaction solution, before adding the selected enzyme. The extent of the reactions was monitored on a time-course basis by sampling the incubation mixture at different time intervals. Proteolytic fragments were analyzed by LC/MS performed on an API 100 single quadrupole electrospray (Applied Biosystems) equipped with two Series 200 LC isocratic pumps (Applied Biosystems). Chromatographic separation was obtained on a C4 column by the means of a 100 single quadrupole LC isocratic pumps (Applied Biosystems) program. Non-deuterated and totally deuterated samples were used as control. The fully deuterated samples were prepared by incubation for 1 h at 25 °C in 50 mM HEPES, pH 7.5, 1 mM DTT, 0.15 mM NaCl, 5 mM GdmHCl in D_2O at 8 μM protein concentration. LC/MS analysis was performed as above.

RESULTS AND DISCUSSION

The M* Intermediate—Previous studies suggested that GdmHCl-induced unfolding at equilibrium of EcAspAT at 20 °C is reversible and proceeds through the formation of at least two monomeric intermediates along the unfolding pathway (33).

\[ N_2 \rightarrow 2M \rightarrow 2U \]

Reaction 1

N2 is the native dimer, M and M* are distinct “structured” monomers, and U is the unfolded state. As described by Herold and Kirschner (33), the first intermediate, M, the folded monomeric EcAspAT, is partially populated at 0.6 M GdmHCl and only at low protein concentrations. The second monomeric intermediate, M*, accumulates at 1 M GdmHCl and constitutes a “molten globule”-like species, which retains part of the tertiary and secondary structure of the native state. The unfolding process of both WT EcAspAT and the P138A mutant was investigated by fluorescence, gel filtration, and CD analysis and the corresponding M* species, identified according to the results reported by Herold and Kirschner (33) and to size-exclusion chromatography data in GdmHCl at equilibrium (34), were thoroughly characterized. Preliminary experiments (25) showed that folding/unfolding equilibrium was attained within 30 min and that, under the experimental conditions used, the folding intermediates were stable. No precipitation of protein was observed after incubation at the protein concentration used, and gel filtration analysis ruled out the presence of soluble aggregates (see below).

Because the enzyme has five tryptophan residues located in various regions of its three-dimensional structure, with one positioned in the active site pocket, tryptophan fluorescence was used as a probe for global changes in tertiary structure of EcAspAT during protein unfolding and to distinguish between M and M* (33). In our experimental conditions, i.e. relatively high protein concentration (in the 1–10 μM range) and the presence of pyridoxal 5'-phosphate, the N2 = 2M transition escaped fluorescence detection. This transition, indeed, is shifted toward higher denaturant concentrations by increasing the protein concentration and by the presence of the cofactor (33), thus overlapping with the M to M* transition.

Fig. 1 shows the biaxochromic shift of the fluorescence maximum for WT EcAspAT and the mutant as the concentration of GdmHCl increased. The unfolding process is clearly biphasic for both proteins, and a stable intermediate state, corresponding to the M* species, could be detected at given GdmHCl concentrations. Therefore, the first fluorescence-detected transition observed included both the dissociation of the dimer and the conversion to M*. Accordingly, the change in the fluorescence spectrum reflects both structural changes and the loss of the coenzyme (PLP) during unfolding.

A clear difference in the emission spectra of the M* intermediates of the two proteins was observed. The fluorescence spectrum of the M* from P138A displayed a maximum at 337 nm, red-shifted by 5 nm as compared with the wild type enzyme (λ_max at 342 nm). This change of the maximum wavelength is significant because no variation greater than ±0.5 nm was...
observed in several spectra recorded for each sample. On the contrary, in the native and in the unfolded states, the emission spectra of P138A and WT were nearly coincident.

Gel filtration analysis of EcAspAT carried out at different GdmHCl concentrations (Fig. 2) showed that the elution volume of the enzyme increased to a broad maximum between 0.7 and 1.2 M GdmHCl and then gradually decreased at higher denaturant concentrations, according to previous reports (33, 34). In all conditions, the protein eluted as a rather symmetric peak, consistent with homogeneous species. Moreover, no differences were detected in the elution volume of WT-M* and P138A-M* at 1.4 M GdmHCl, indicating that no differences in the compactness of the intermediates could be inferred by these analyses. However, both M* forms showed an elution volume in between that of the native N2 and that expected for the folded monomer M, thus indicating that in these conditions the M* species are monomeric and have a compact structure, albeit not as structured as the folded monomeric protein. Therefore, all the experiments described below were carried out at equal or lower EcAspAT concentrations to ensure the monomeric state of the protein.

The native, M*, and denatured states of the two proteins were then analyzed by circular dichroism. Fig. 3 shows the CD spectra of the three forms of P138A, which were coincident with those recorded for the respective WT forms (data not shown). In both cases, the M* species showed a CD signal at 222 nm that is 58% of the native protein, indicating that these intermediates have a considerable amount of apparent secondary structure.

**Tryptophan Fluorescence Quenching**—The native and M* states of the two proteins were then analyzed by circular dichroism. Fig. 3 shows the CD spectra of the three forms of P138A, which were coincident with those recorded for the respective WT forms (data not shown). In both cases, the M* species showed a CD signal at 222 nm that is 58% of the native protein, indicating that these intermediates have a considerable amount of apparent secondary structure.

**Tryptophan Fluorescence Quenching**—The native and M* states of the two proteins were then analyzed by circular dichroism. Fig. 3 shows the CD spectra of the three forms of P138A, which were coincident with those recorded for the respective WT forms (data not shown). In both cases, the M* species showed a CD signal at 222 nm that is 58% of the native protein, indicating that these intermediates have a considerable amount of apparent secondary structure.

**Tryptophan Fluorescence Quenching**—The native and M* states of the two proteins were then analyzed by circular dichroism. Fig. 3 shows the CD spectra of the three forms of P138A, which were coincident with those recorded for the respective WT forms (data not shown). In both cases, the M* species showed a CD signal at 222 nm that is 58% of the native protein, indicating that these intermediates have a considerable amount of apparent secondary structure.

The national, M*, and denatured states of the two proteins were then analyzed by circular dichroism. Fig. 3 shows the CD spectra of the three forms of P138A, which were coincident with those recorded for the respective WT forms (data not shown). In both cases, the M* species showed a CD signal at 222 nm that is 58% of the native protein, indicating that these intermediates have a considerable amount of apparent secondary structure.

Acrylamide quenching experiments on the M* intermediates were impaired by protein aggregation occurring even at low acrylamide concentrations. In contrast, quenching experiments with iodide could be carried out on the M* intermediates at 0.39 μM protein concentration without major problems. The results reported in Fig. 4B demonstrated that the tryptophan residues responsible for the fluorescence of the M* intermediate are less exposed in P138A than in WT, as also shown by the
**Table I**

| Protein | \(K_q \) native state \(a\) | \(K_q \) M* state \(a\) |
|---------|-------------------|-------------------|
| WT      | 0.718 ± 0.078     | 5.630 ± 0.052    |
| P138A   | 0.759 ± 0.036     | 4.533 ± 0.055    |

\(^a\) \(K_q \) values were determined with Equation 3.

\(^b\) \(K_q \) values were determined with Equation 1.

**K** \(_q\) values reported in Table I. Iodide quenching data were analyzed to determine the number of tryptophan residues involved in quenching. From the \(f_a\) values calculated by Equation 2, 4.30 (± 0.30) of 5 total tryptophan residues could be quenched in the M* intermediate of WT as compared with only 3.05 (± 0.25) in P138A. The quenching data obtained for both WT and P138A in the denatured state were interpolated with the same equation as a control, giving an \(f_a\) value of 1, thus indicating that all tryptophans are exposed in the denatured form (data not shown). Again these data suggested a different arrangement of the three-dimensional structure in the M* intermediates from WT and P138A.

Equilibrium and Kinetic Intermediates—Fig. 5A shows the kinetic process for the refolding of WT at 5 °C from fully unfolded state (5 mM GdmHCl) by 50-fold dilution into refolding buffer (0.1 mM K-HEPES, pH 7.5, 1 mM DTT), at a final protein concentration of 0.39 mM. A transient intermediate \(I_1^*\) accumulated within 50 s and then slowly evolved to the native state in a biphasic curve that can be fitted to a double exponential with the kinetic constants reported in Table II. The folding events leading to \(I_1^*\) escaped detection; however, using the extrinsic probe for hydrophobic surfaces, ANS, the formation of a further folding intermediate \(I_2^*\) within the dead time of manual mixing was observed (Fig. 5B). It is worth mentioning that ANS does not bind either the unfolded EcAspATs or the native proteins, whereas the M* form of both WT and P138A interacts with ANS, as shown by increased fluorescence intensity emission at 475 nm upon excitation at 375 nm of the extrinsic probe (data not shown).

When the formation of M* was initiated by diluting the WT protein from 5 mM GdmHCl to 1.4 mM GdmHCl and monitored by tryptophanyl fluorescence (Fig. 5C), a constant signal was obtained within the dead time of manual mixing. Following a further GdmHCl dilution to the native state, the fluorescence signal rapidly increased up to an intensity value similar to that of the \(I_1^*\) detected in the refolding experiment described above (see Fig. 5A) and then slowly decreased to the native value. Therefore, in WT the equilibrium M* intermediate is clearly distinct from the transient intermediate \(I_1^*\) that could still be detected in the refolding pathway from M* to N\(_2\). When the same experiment was carried out on P138A, a different behavior was observed (Fig. 5D). Formation of M* from the unfolded mutant occurred with a slower kinetic than in the WT, and the M* intermediate was characterized by the same tryptophanyl fluorescence intensity at 333 nm displayed by the \(I_1^*\) species. Moreover, following dilution to the native state, no increase in the fluorescence intensity was detected with the signal simply decaying to the \(N_2\) value, indicating that no other intermediate accumulated on the refolding pathway from M* to \(N_2\). Therefore in P138A the spectroscopic features of the M* intermediate are very similar to those of the \(I_1^*\) transient species, suggesting that M* might be an on-pathway folding intermediate possibly coincident with \(I_1^*\).

These results confirmed the hypothesis that cis-prolines play a subtle role in directing the traffic of intermediates toward the unique structure of the native state, rather than to respond to the needs for specific catalytic or functional roles (25). However, the involvement of trans-cis isomerization of the peptide bond preceding Pro-138 is not as simple as hypothesized by Leistler and colleagues (35), who indicated this event as one of the possible rate-limiting steps leading to the two slow phases observed in WT refolding. Addition of a peptidyl prolyl isomerase, cyclophilin A, to the refolding solution did not alter either the fluorescence decay or the activity recovery (data not shown). Nevertheless, the ineffectiveness of cyclophilin A to catalyze the two slow phases, does not rule out an involvement of isomerization in the refolding of EcAspAT either at the early stages of the process or when the trans peptide bond is already buried in the protein and is then not accessible to the isomerase. It should be noted that refolding of P138A from 5 mM GdmHCl still shows the presence of two slow phases with kinetic constants similar to those calculated for WT (Table II), even in the absence of any cis-trans isomerization at Ala-138.

**Hydrogen/Deuterium Exchange—**Isotopic exchange at peptide bond protons basically depends on whether they are participating in intramolecular hydrogen bonding and on the extent of solvent shield exerted by the protein structure. Measurements of protein mass increase following hydrogen/deuterium exchange may then be used as a sensitive probe to monitor conformational changes in protein structure. A comparative characterization of \(N_2\) and M* species from WT and P138A was carried out by hydrogen/deuterium exchange experiments followed by ES/MS analysis. Protein solutions (20 μM) were 10-fold diluted by the addition of the appropriate D\(_2\)O buffers, and deuterium incorporation was monitored by sampling the incubation mixture at different interval times followed by cold acqueous and fast LC/MS analysis.

Fig. 6A shows the number of amide protons exchanged with deuterium in the native form of WT and P138A as a function of time. Following 60 min of reaction, a total of 120 ± 6 hydrogens were replaced with deuterium atoms in both proteins. This value corresponds to ~30% of the theoretically exchangeable protons, thus confirming the high level of compact structure of the native form of both proteins. However, a clear difference in the exchange kinetics between the WT and the mutant was detected. Incorporation of deuterium atoms occurred much faster in P138A than in WT, suggesting a higher degree of flexibility for the \(N_2\) state of the mutant.

The CD and fluorescence studies of the native state of the two proteins reported above had yielded almost indistinguishable spectroscopic characteristics. Moreover, inspection of x-ray data indicated that the crystallographic structures of P138A and WT are superimposable but for a small region surrounding the mutation site (25). Nevertheless, some differences in the conformation of the two proteins could be detected by H/D exchange experiments. These differences should very likely be related to the dynamics of the proteins and are then not readily observed in static three-dimensional structures. The faster kinetics of deuterium incorporation observed in P138A could in fact be interpreted in terms of a wider protein breathing experienced by the mutant with respect to WT. On the other hand, the same extent of total proton exchange shown by the two proteins confirms their nearly identical three-dimensional structure.

The exchange experiments on the M* species were performed by incubation of the intermediates in deuterated buffer for 60 min followed by ES/MS analysis, from which a number of considerations could be drawn. The single envelope of isotope peaks occurring in the mass spectra (Fig. 6B) indicated an EX1 kinetics of hydrogen and deuterium exchange according to Bai et al. (32), confirming that the M* species of both P138A and WT are homogeneous and stable folding intermediates. The isotope pattern distribution is an important source of information on sample heterogeneity. For structurally heterogeneous
samples, such as the mixture of folded and unfolded molecules, or the presence of different aggregation states, multiple envelopes of isotope peaks are expected, each with different hydrogen exchanging characteristics, whereas homogeneous samples show a single envelope of isotope peaks at all incubation times.

As expected, a much higher extent of isotopic exchange was observed for the M* species with respect to the native forms; M* intermediates from P138A and WT incorporated 245 ± 7.8 and 286 ± 9 deuterium atoms corresponding to 61 and 71% of the exchangeable protons, respectively. The large increase in the number of exchanged protons might be ascribed both to the dissociation of the dimeric native state of the proteins into the monomeric forms and to the partial unfolding of the protein structure leading to the M* state. Because 40% of the amide protons in P138A and more than 32% in WT are still shielded to the solvent in the M* form, these results confirm previous spectroscopic data that the M* species are partly folded intermediates still retaining a significant amount of secondary and tertiary structure.

Finally, the M* from P138A exchanged significantly fewer hydrogen atoms than the corresponding intermediate from WT, ~40 protons corresponding to nearly 16% of the total exchanged hydrogens. This result demonstrated that the M* intermediates from P138A and WT are different, confirming the spectroscopic data reported above, with the M* species from P138A exhibiting a more compact and less flexible structure.

**Limited Proteolysis of M* Intermediates from P138A and WT**—The surface topology of the M* intermediates from P138A and WT was investigated by a strategy that combines limited proteolytic digestions with mass spectrometric identification of the released fragments. Limited proteolysis experiments were performed using trypsin, chymotrypsin, endoprotease Lys C, and subtilisin as proteolytic probes under conditions in which the M* intermediates were monomeric and exhibited maximal stability. The M* species were incubated with each protease using an appropriate enzyme to substrate ratio, and the extent of the enzymatic hydrolysis was monitored on a time-course basis by sampling the incubation mixture at different interval times followed by LC/MS analysis. The chromatographic and mass spectrometric identification of the fragments released from the protein led to the assignment of preferential cleavage sites.

![Fig. 5. Refolding of EcAspAT followed by fluorescence.](image)

**TABLE II**

Kinetic constants calculated for the two slow phases of refolding of WT and P138A from intrinsic fluorescence intensity decay

|       | $k_1$       | $k_2$       |
|-------|-------------|-------------|
| WT    | $2.139 \times 10^{-3} \pm 6.4 \times 10^{-5}$ | $5.05 \times 10^{-4} \pm 1.3 \times 10^{-5}$ |
| P138A | $1.880 \times 10^{-3} \pm 6.4 \times 10^{-6}$ | $1.95 \times 10^{-4} \pm 6.8 \times 10^{-6}$ |
Fig. 6. Hydrogen/deuterium exchange in EcAspAT monitored by ES/MS. A, the number of exchanged protons in WT-N$_2$ (■) and P138A-N$_2$ (●) was calculated by the increase in the molecular mass of the two proteins and is reported as a function of the exchange time. B, electrospray m/z signals of the +40 and +34 ions of P138A-M$^+$ (a) and WT-M$^+$ (b) after 60 min of incubation in buffered D$_2$O.
The preferential cleavage sites observed in M* from P138A essentially gathered into two separate regions of the protein, i.e. the segments 98–121 and 266–288, whereas few isolated sites were detected at Tyr-59, Ser-136, Ala-150, and Tyr-380. It should be noted that no cleavage was detected in the N-terminal arm that in the dimeric form of EcAspAT embraces the partner subunit concurring in the formation of the active site. This observation suggests that, following dissociation of N2 into M, the monomeric form should undergo conformational changes to generate the M* species in which the N-terminal tail is forced to interact with the protein body and is then protected from the proteases. Otherwise, the flexible N-terminal arm would have been cleaved by proteases at several sites.

A greater number of cleavage sites and a faster kinetic of hydrolysis was observed in all the experiments on the M* species from WT, confirming that the two intermediates exhibit a different three-dimensional arrangement with the conformation of M* from WT being generally more flexible and less structured than P138A-M*. These results are in excellent agreement with previous H/D exchange experiments depicting a higher flexibility of the WT-M* species that showed a larger extent of deuterium incorporation.

Accordingly, besides the exposed regions already observed in P138A-M*, additional proteolytic sites in WT-M* were identified within the segments 20–25 and 39–41 with few isolated cleavages observed at Tyr-59, Tyr-161, and Phe-188. Particularly, the region 39–41 and Tyr-161 are located at the inter-domain interface. Occurrence of proteolytic cleavages at this region only in WT-M* suggests that the structural differences between the two M* intermediates might be ascribed to a different relative orientation of the two domains. The relative rotation of the two domains might also affect the conformation of the N-terminal tail, thus explaining the accessibility of the segment 20–25, located in the small domain close to the N-terminus of the protein.

Comparative proteolysis experiments were also performed on the native dimeric forms of P138A and WT, using the same conditions employed for the M* intermediates. Both proteins were completely resistant to proteases, and no cleavage was detected at any time of incubation even when the E/S ratio was increased up to 50-fold.

Conclusion—A combination of spectroscopic techniques, H/D exchange, and limited proteolysis experiments coupled to mass spectrometry analysis were used to depict the topology of the M* partly folded intermediate of EcAspAT in WT as well as in the mutant protein P138A.

The role of intermediates in the protein folding is a controversial issue. In some cases, they appear to be important milestones for productive folding (i.e. they are on-pathway), whereas, in other cases, they might arise by nonspecific collapse of the polypeptide chain or accumulate because they are trapped by non-native interactions (i.e. they are off-pathway). M* is an off-pathway intermediate in the folding process of WT EcAspAT, as already pointed out by Leistler et al. (35), whereas spectroscopic data provided in this study suggest that the M* species from the mutant P138A is coincident with the on-pathway folding intermediate I1*.

Gel filtration and CD analysis did not show major differences between the two M* species. On the contrary, the increased deuterium incorporation level and the higher accessibility to proteases shown by WT-M* as compared with P138A-M* demonstrated the occurrence of conformational differences between the M* intermediates. These results together with fluorescence data indicate that P138A-M* is conceivably more compact than WT-M*.

In particular, limited proteolysis results suggested that
these conformational differences might be related to a different relative orientation of the small and large domains. These data can be compared with those reported by Martinez-Carrion and co-workers (36, 37) during the early stages of refolding of mitochondrial AAT (mAspAT) that shares 42% identity with EcAspAT. Although mAspAT does not show stable folding intermediates, the pattern of preferential proteolytic sites in the N-terminal region are in good agreement with that obtained on WT-M*. Because cis-Pro-138 is conserved in mAspAT, this common behavior might indicate that isomerization of the 137–138 peptide bond during the folding process affects the relative rotation of the two domains causing the exposition of the N-terminal region. These conformational changes do not occur in the P138A mutant, where the 137–138 peptide bond is in the trans configuration, originating a more compact structure and preventing proteolysis in the N-terminal region.

In the light of the conformational differences between the two M* species, it is clear that the peptide bond preceding Pro-138 in WT-M* is in cis conformation at this stage of the folding process. A direct link between cis-trans isomerization at position 138 and folding of EcAspAT cannot be easily depicted (25). This event certainly plays some role, as demonstrated by the different characteristics of the two M*, and, even more, by the slower reactivation of P138A (25). However, other rearrangements related to the cis-trans isomerization may rule the folding process of this protein. For instance, some favorable tertiary interactions in key folding intermediates might be formed only following cis-trans isomerization at position 138, thus leading to a slower reactivation in P138A rather than in WT, or, alternatively, isomerization at position 138 occurring at the early stages of refolding, might direct the process into a different route.

Finally, it is intriguing to observe that the H/D exchange results showed few differences in the native N₂ forms of WT

---

**TABLE III**

| Preferential cleavage sites detected on the P138A-M* and WT-M* intermediates in the different limited proteolysis experiments |
|---------------------------------------------------------------|
| **Differently exposed residues are highlighted in bold type.** |
| **P138A-M** | **WT-M** |
| Trypsin | Arg-101, Arg-266 | Arg-25, Arg-101, Arg-266 |
| Chymotrypsin | Tyr-59, Tyr-380 | Phe-24, Tyr-40, Tyr-59, Tyr-161, Phe-188, Tyr-256, Tyr-380 |
| Lys C | Lys-98, Lys-121, Lys-288 | Lys-41, Lys-98, Lys-121, Lys-288 |
| Subtilisin | Ala-115, Ser-136, Ala-150, Cys-270, Tyr-380 | Leu-20, Val-39, Phe-188, Cys-270, Tyr-380 |

---

**Fig. 8. Map of cleavage sites on EcAspAT.** Preferential proteolytic sites are mapped onto the amino acid sequence of EcAspAT. Cleavage sites only occurring in WT-M* are indicated by dashed arrows, those only present in P138A-M* by empty arrows, and those observed in both proteins by filled arrows. Dotted, underlined, and double underlined regions indicate N-terminal tail, the large domain, and the small domain, respectively. Residues occurring at the domain interface are highlighted in gray. Amino acid numbering is based on the sequence of porcine cytosolic aspartate aminotransferase (38).
A Partly Folded Intermediate of EcAspAT

and P138A, the spectroscopic features and crystallographic structures of which are almost superimposable. Flipping of the 137–138 peptide bond from the cis conformation of WT to the trans conformation of P138A resulted in an overall increased "motility" of the fraction of the protein that is accessible to the solvent in the native form.

Limited proteolysis and H/D exchange experiments in conjunction with mass spectrometry analysis well complemented spectroscopic results, thus providing a wealth of data that allowed a detailed structural analysis of EcAspAT folding intermediates. On a more general ground, these results suggest that the integration of spectroscopic investigations and mass spectrometric procedures might be instrumental in providing subtle structural details on transient conformations such as those present along a folding pathway.

Acknowledgment—We thank Dr. Alessia Errico for preliminary limited proteolysis experiments.

REFERENCES

1. Dobson, C. M., and Karplus, M. (1999) Curr. Opin. Struct. Biol. 9, 92–101
2. Chiti, F., Tabhei, N., Buciantini, M., White, P., Ramponi, G., and Dobson, C. M. (2000) EMBO J. 3, 1441–1449
3. Englander, S. W., and Kalchenbach, N. R. (1984) Q. Rev. Biophys. 16, 521–655
4. Li, R., and Woodward, C. (1999) Protein Sci. 8, 1571–1590
5. Woodward, C., Simon, I., and Tuchsen, E. (1982) Mol. Cell. Biochem. 48, 135–160
6. Johnson, R. S., and Walsh, K. A. (1994) Protein Sci. 3, 2411–2418
7. Smith, D. L., Deng, Y., and Zhang, Z. (1997) J. Mol. Biol. 268, 223–239
8. Ratta, V., and Chari, B. T. (1995) Rapid Commun. Mass Spectrom. 5, 214–217
9. Zhang, Z., and Smith, D. L. (1993) Protein Sci. 2, 522–531
10. Miranker, A., Robinson, C. V., Aplin, R. T., and Dobson, C. M. (1999) Science 282, 896–899
11. Zhang, Z., and Smith, D. L. (1996) Protein Sci. 5, 1282–1289
12. Engen, J. R., Greiner, W. H., Smithgall, T. E., and Smith, D. L. (1999) Biochemistry 38, 8926–8935
13. Halgand, F., Dumas, R., Biou, V., Andrieu, C., Thomazeau, K., Gagnon, J., Douce, R., and Forest, E. (1999) Biochemistry 38, 6025–6034
14. Zappacosta, F., Pessi, A., Bianchi, E., Venturini, S., Sollazzo, M., Tramontano, A., Marino, G., and Pucci, P. (1996) Protein Sci. 5, 882–883
15. Scognamiglio, R., Notomista, E., Barbieri, P., Pucci, P., Dal Piaz, F., Tramontano, A., and Di Donato, A. (2001) Protein Sci. 10, 482–490
16. Orrù, S., Dal Piaz, F., Casharba, A., Biasiol, G., De Francesco, R., Steinkühler, C., and Pucci, P. (1999) Protein Sci. 8, 1445–1454
17. Bianchi, E., Orrù, S., Dal Piaz, F., Ingenito, R., Casharba, A., Biasiol, G., Koch, U., Pucci, P., and Pessi, A. (1999) Biochemistry 38, 13844–13852
18. Urbani, A., Biasiol, G., Brunetti, M., Volpapi, C., Di Marco, S., Sollazzo, M., Orrù, S., Dal Piaz, F., Casharba, A., Pucci, P., Nardi, C., Gallinarini, P., De Francesco, R., and Steinkühler, C. (1999) Biochemistry 38, 5206–5215
19. Esposito, G., Michelutti, R., Verdone, G., Viglino, P., Hernandez, H., Robinson, C. V., Amoresano, A., Dal Piaz, F., Monti, M., Pucci, P., Mangione, P., Stoppini, M., Merlina, G., Ferrari, G., and Bellotti, V. (2000) Protein Sci. 9, 831–845
20. De Lorenzo, C., Dal Piaz, F., Piccoli, R., Di Marco, A., Pucci, P., and D’Alessio, G. (1998) Protein Sci. 7, 2653–2658
21. Piccoli, R., De Lorenzo, C., Dal Piaz, F., Pucci, P., and D’Alessio, G. (2000) J. Biol. Chem. 275, 8000–8006
22. Scaloni, A., Miraglia, N., Orrù, S., Amodeo, P., Motta, A., Marino, G., and Pucci, P. (1999) J. Mol. Biol. 277, 945–958
23. Scaloni, A., Monti, M., Acquaviva, R., Tell, G., Damante, G., Formisano, S., and Pucci, P. (1999) Biochemistry 38, 64–72
24. Atkinson, R. A., Joseph, C., Dal Piaz, F., Biolo, L., Steier, G., Pucci, P., and Pastore, A. (2000) Biochemistry 39, 5255–5264
25. Biolo, L., Malashkevich, V. N., Capitani, G., De Luca, F., Moretta, A., Jansons, J. N., and Marino, G. (1999) Biochemistry 38, 905–913
26. Jager, J., Moser, M., Sauter, U., and Janssen, J. N. (1996) J. Mol. Biol. 239, 285–305
27. McPhalen, C. A., Vincent, M. G., Fibot, D., Jansons, J. N., Lesk, A. M., and Chothia, C. (1992) J. Mol. Biol. 222, 197–213
28. Yano, T., Kuramitsu, S., Tanase, S., Morino, Y., Hiromi, K., and Kagamiyama, H. (1991) J. Biol. Chem. 266, 6079–6085
29. Kuramitsu, S., Hiromi, K., Hayashi, H., Morino, Y., and Kagamiyama, H. (1990) Biochemistry 29, 5469–5476
30. Lehrer, S. S. (1971) Biochemistry 10, 3254–3263
31. Efrink, R. M., and Ghiron, C. A. (1976) Biochemistry 15, 672–680
32. Bai Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Proteins 17, 75–86
33. Herold, M., and Kirshner, K. (1990) Biochemistry 29, 1907–1913
34. Herold, M., and Leistler, B. (1992) FEBS Lett. 288, 26–29
35. Leistler, B., Herold, M., and Kirschner, K. (1992) Eur. J. Biochem. 205, 603–611
36. Mattingly, J. R., Jr., Torella, C., Iriarte, A., and Martínez-Carrion, M. (1998) J. Biol. Chem. 273, 23191–23202
37. Torella, C., Mattingly, J. R., Jr., Artigues, A., Iriarte, A., and Martínez-Carrion, M. (1998) J. Biol. Chem. 273, 3915–3925
38. Ovcinnikov, Y. A., Egorov, C. A., Aldanova, N. A., Feigina, M. Y., Lipkin, V. M., Abdulrae, N. G., Grishin, E. V., Kiselev, A. P., Modyanov, N. N., Braunstein, A. E., Polyanovsky, O. L., and Nosikov, V. V. (1973) FEBS Lett. 29, 31–34