シロイヌナズナにおけるコ-シャペロニンの解析

河本　恭子

要旨

生物の持つ遺伝情報は DNA→mRNA→タンパク質というセントラルドグマに従い伝達され、機能を発揮する。DNA、RNA という核酸は主としてその 1 次構造が重要であるのに対して、タンパク質は正しい高次構造をとる必要がある。高次構造形成（巻き戻り）は遺伝情報発現における重要な過程の 1 つであると考えられる。巻き戻りは自発的におこると考えられてきたが、近年いわゆる分子シャペロンと呼ばれる一群のタンパク質が重要な機能を果たすことが明らかになってきた。分子シャペロンの中でもシャペロニンシステムはタンパク質の巻き戻り過程に直接関与する重要な因子である。そこで、本研究ではシャペロニンシステムに注目し、高等植物においてその全体像の解明を目的とした。

シャペロニンシステムは原核細胞から真核細胞まで全ての生物が共通に持っている。シャペロニンシステムはシャペロニン（Cpn）とコ-シャペロニンから成る。最も研究の進んでいるのは大腸菌のシャペロニン（GroEL）である。GroEL は 57kDa のタンパク質で、その 14 量体が形成する筒状の構造はタンパク質の巻き戻りのための‘場’となる。GroEL は GroES と呼ばれる 10 kDa の調節因子を必要とし、このような調節因子はコ-シャペロニンと総称されている。本研究においてはまず、コ-シャペロニンについて解析を行った。

シロイヌナズナのコ-シャペロニンホモログを同定するため、大腸菌の GroES 温度感受性突然変異体を相補する cDNA クローンを単離した。得られたクローンは 10 kDa のコーシャペロニン（Cpn10）ホモログをコードしていた。動物のミトコンドリアの Cpn10 に 50%程度の高い同一性を示したことからミトコンドリア Cpn10 のホモログ（AT1）

であると予想された。Cpn10（AT1）はシロイヌナズナのロゼット葉を用いた細胞分画や、過剰発現する形質転換タバコを用いた免疫電顕により、ミトコンドリアに局在すること確かめられた。

葉緑体のコ-シャペロニンホモログである Cpn20 は Cpn10（GroES）様のドメインが 2 つタンデムにつながった構造をもつ。シロイヌナズナの Cpn20 の cDNA クローンを単離した結果、予想される成熟タンパク質部分はホウレンソウのものと 61%の同一性を示した。過剰発現する形質転換タバコを用いた免疫電顕により、葉緑体に局在することが確かめられた。GroES は 7 量体で機能することが知られているが、Cpn20（AT2）の多量体構造を調べた結果から 4 量体を形成することがわかった。Cpn20（AT2）の 4 量体には GroES 様ドメインが 8 個含まれることになる。Cpn20（AT2）は GroES と同様に、GroEL と相互作用しコ-シャペロニンとして機能できる。GroEL と Cpn20（AT2）の複合体に含まれるそれぞれの量比は 14 : 3.6 であったので、GroEL の 14 量体と Cpn20（AT2）の 4 量体が結合していることが強く示唆された。GroES はモバイルループと呼ばれる突出した部位で GroEL と結合する。GroES と GroEL の複合体では、1 つのモバイルループが 1 つの GroEL と結合していると考えられる。一方、Cpn20（AT2）の多量体は 8 カ所のモバイルループを持つため、GroEL との結合は 1 対 1 ではなく不規則で柔軟な相互作用をしていると考えられる。

さらに、EST データバンクを検索することにより、アミノ末端に延長配列を持つ Cpn10 ホモログ（AT3）が見つかった。既知のコ-シャペロニンホモログの系統樹を作成したところ、Cpn10（AT3）はどのホモログとも相同性が低いことがわかった。結晶構造が解析されている GroES との比較から、Cpn10（AT3）はルーフ構造に当たる領域が大きく変化したコ-シャペロニンであると推測された。T4 ファージのコ-シャペロニンホモ

ログである Gp31 は、より大きな基質に対応するためにループ構造を欠失する形に進化したと考えられている。Cpn10（AT3）もループ構造に変異が変化していることから、特異的な基質が存在する可能性がある。Cpn10（AT3）はチラコイド内腔に存在する可能性があり、葉緑体の発達と関連した発現パターンを示すことからもチラコイド膜に多く含まれる光合成に関わるタンパク質が基質である可能性が考えられた。

本研究では、シロイヌナズナにおいて 3 種類のコ-シャペロニンホモログを同定した。その結果、ミトコンドリアのシャペロニンシステムは大腸菌の GroEL-GroES 型とよく似ていることがわかった。一方葉緑体には複数のシャペロニンシステムが存在し、GroEL-GroES 型とは異なる特徴を持っていた。おそらく基質となるタンパク質の特異性に合わせて変化してきたのだろうと考えている。

目次

略号

BSA: bovine serum albumin

CCT: chaperonin containing TCP1

Cpn: chaperonin

CS: citrate synthase

DSP: 3, 3'-dithiobis (succinimidyl propionate)

Hsp70: heat shock protein 70

MALDI-TOF: matrix-assisted laser desorption ionization-time-of-flight

RuBisCO: ribulose 1,5-bisphosphate carboxylase/oxygenase

PAGE: polyacrylamide gel electrophoresis

PCR: polymerase chain reaction

TCP1: t-complex containing polypeptide1

TF55: thermophilic factor 55

序論

分子シャペロン

生物の持つ遺伝情報は DNA→mRNA→タンパク質というセントラルドグマに従い伝達され、機能を発揮する。DNA、RNA という核酸は主としてその 1 次構造が重要であるのに対して、タンパク質は正しく巻き戻る必要がある。最近の研究では、正しく巻き戻れないタンパク質はそれ自体の機能を失うだけでなく、細胞全体にストレスを与えることが示されている。例えば、異常タンパク質の蓄積がアルツハイマー病などの神経疾患の原因であることが示されてきている。このようにタンパク質が正しい高次構造をとることは生命活動の維持に必須であり、その機構を理解することは非常に重要である。

転写、翻訳の過程は核酸が相補的な塩基対を形成することにより、機械的に正確な反応が起こる。一方、タンパク質の巻き戻り（高次構造形成）の過程は、タンパク質のアミノ酸配列によってかなり異なっている。In vitro でタンパク質の高次構造を一旦破壊した後、希釈により変性剤濃度を低くすると巻き戻りが観察できるものがある。このことから巻き戻りは自発的におこると考えられた（Anfinsen、1973)。しかしながら、in vitro で巻き戻りが観察される場合でも巻き戻りの効率を上げるためには巻き戻りの条件を検討する必要がある。タンパク質の濃度や反応温度を低くすることは、分子間の凝集を防ぎ分子内での巻き戻りを促進させる効果がある。in vitro で効率よい巻き戻りに必要なこれらの条件は、in vivo の状況を反映したものとはいえず、生体内では何らかの機構が必

要であろうと考えられていた。

細胞内での巻き戻りには、いわゆる分子シャペロンと呼ばれる一群のタンパク質が必要な場合があることが明らかにされてきた（Feldman と Frydman、2000)。分子シャペロンはタンパク質の巻き戻り、会合、膜の通過などに関与するが、最終的な構造体には含まれない一群のタンパク質に対して付けられた名前である（Ellis と van der Vies、1991)。分子シャペロンは当初、熱ストレスで誘導されるタンパク質（heat-shock protein, HSP）として注目された。熱により変性したタンパク質に結合して細胞への悪影響を抑えたり、変性したタンパク質の再生に働くと考えられた。その後、必ずしも熱ストレスで誘導されるものばかりではないこと、また、誘導されるものであっても通常の生育温度においても必須であることがわかってきた（Fayet ら、1989)。

分子シャペロンは真核、原核細胞に共通して存在しており、代表的なものにはシャペロニンシステムと Hsp70 がある。シャペロニンシステムは“Anfinsen cage”とも呼ばれ、新生タンパク質がその中で正しく巻き戻るための‘カゴ’となる。この‘カゴ’は、in vitro の巻き戻り実験において試行錯誤していた分子間の凝集の阻止という問題を見事に解決するものであった。Hsp70 は翻訳中のポリペプチド鎖に結合し、ほどけた状態を安定に保持したままシャペロニンシステムへ受け渡すという役割を持つ。このように、シャペロニンシステムはタンパク質の高次構造形成という遺伝情報発現における重要な過程に、直接関与する分子である。そこで私は分子シャペロンの中でも特にシャペロニンシステムに注目し、高等植物におけるシャペロニンシステムの全体像の解明を目的とした。

シャペロニンシステムの機能

シャペロニンシステム

シャペロニン（Cpn）は1次配列上からは2つのグループに分類される（Braig、1998、図1）。グループ1は大腸菌などの真生細菌と真核細胞のミトコンドリアや葉緑体、グループ2は古細菌と真核細胞の細胞質という進化上関連深い場所に存在する（Braig、1998）。両者は機能的に同等ながら1次配列や多量体構造などはかなり異なっており、オルガネラが共生した真正細菌に由来するという細胞内共生説を裏付けるものとなっている。グループ1はシャペロニン（Cpn60）の働きを制御するコ-シャペロニンが存在するのに対し、グループ2はコ-シャペロニンを必要としない。グループ2は、真核生物の細胞質に存在するCCT（chaperonin containing TCP1）、古細菌の持つサーモソームやTF55（thermophilic factor 55）を含む。他のシャペロニンと特に異なるCCTの特徴としては、それぞれ相同性の低いポリペプチド鎖より多量体を形成するという点が挙げられる。

グループ1のシャペロニンシステム

最も研究の進んでいるのは大腸菌のシャペロニン（GroEL）、コ-シャペロニン（GroES）である。GroEL、GroESの多量体構造（Braigら、1994、Huntら、1996）、GroEL –GroES複合体の構造（Xuら、1997）が既に結晶構造解析されており、シャペロニンシステムの作り出す“巻き戻りの場”は原子レベルで詳細に調べられている。GroELは57 kDaのタンパク質で、7量体が頭と頭をつけた形で2層に積み重なった14量体として働く（Braigら、1994）。GroELの14量体が形

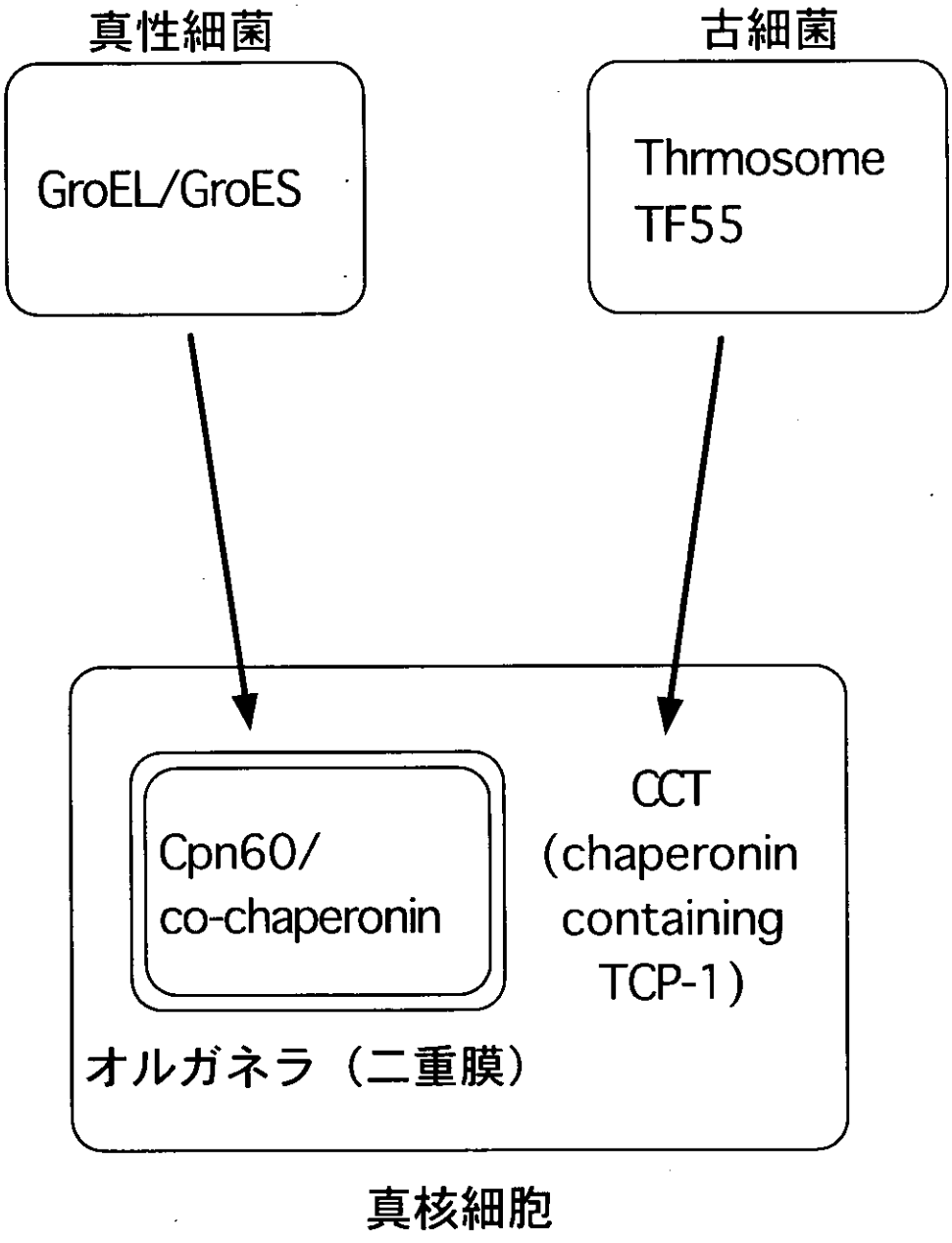


図１　シャペロニンシステムの局在を示すモデル図。
２種類のシャペロニンシステムは進化上関連深い場所で機能している。

成する構造は筒状で、中央に大きな穴を持つ。この中央の穴にこれから巻き戻りを行おうとするほどけたポリペプチド鎖が結合する（図 2)。ほどけたポリペプチド鎖は本来タンパク質内部に存在する疎水性の領域が露出し、凝集しやすい性質を持つ。そのため、Hsp70 などの他の分子シャペロンが一時的に結合して、GroEL へ受け渡すと考えられている。一方、10 kDa の小さなタンパク質である GroES は 7 量体のドーム上の構造をとり、GroEL の 14 量体のどちらかの端に結合する。ポリペプチド鎖と同じ側に GroES が結合すると、ちょうど穴にふたをしてポリペプチド鎖を閉じこめたようになる。GroES の結合により GroEL の 14 量体の構造は大きく変化し、GroEL の穴の内部は疎水的な環境から親水的な環境に変化する（Braig ら、1994)。このため、ポリペプチド鎖と GroEL の結合は弱まり、ポリペプチド鎖の分子内での巻き戻りが促進されるのではないかと考えられている。さらに、GroES の結合により生じた構造変化は何も結合していない反対側にも影響を与え、7 個の GroEL の ATP 加水分解を協調的に引き起こす。この ATP 加水分解の結果、GroES と巻き戻ったタンパク質が GroEL より放出される。巻き戻りにくいポリペプチド鎖は一度のサイクルでは完全に巻き戻らず、再度 GroEL に取り込まれるというサイクルを繰り返して機能を持ったタンパク質に巻き戻るということが示されている（Weissman ら、1995)。

シャペロニンシステムの in vivo での解析

基本的なシャペロニンシステムの作用機構は明らかにされたが、in vivo での機能についてはわからないことも多い。Horwich ら（1993）は大腸菌の GroEL に変異を導入し、温度感受性の致死突然変異株を作製した。この変異株は非許容

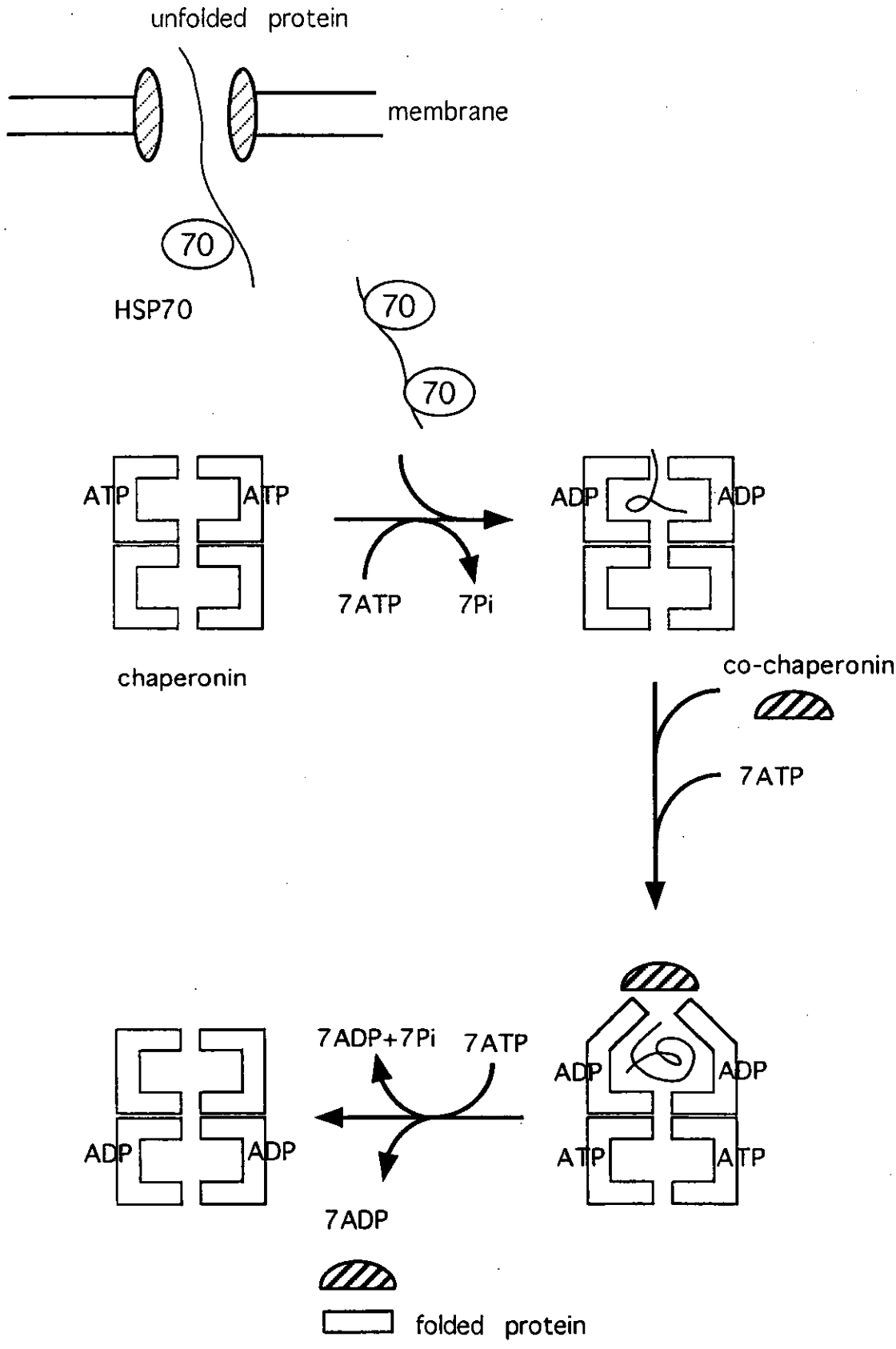


図2　シャペロニンの関与するタンパク質の巻き戻り過程を示すモデル図。
各タンパク質の名称は図中に示した。

温度に移すと、約 1 時間後から生育速度が遅れ、致死となる。温度を上げたときに新しく翻訳されたポリペプチド鎖を 2 次元電気泳動で解析すると、野生型では可溶性であるのに変異株では凝集してしまうタンパク質が生じていた。凝集してしまうタンパク質は巻き戻る際に GroEL を必要とすると考えられる。このような解析の結果、可溶性タンパク質の約 30%の巻き戻りに GroEL が関与していることが示唆された。また、大腸菌を［^{35}S］メチオニンでパルスラベル後、GroEL の抗体で免疫沈降させた結果、ラベルされたタンパク質のうち 10～15%が GroEL と共沈した（Ewalt ら、1997）。これらの結果は大腸菌のほとんどのタンパク質はシャペロニンシステムに依存せず巻き戻ることができるという意外なものであった。

タンパク質の性質がそれぞれ大きく異なる様に、シャペロニンシステムへの依存性にもタンパク質による差があると考えられる。酵母のミトコンドリアのシャペロニンシステムはシャペロニン（Cpn60）とコ-シャペロニン（Cpn10）から成り立っているが、それぞれの温度感受性突然変異株、及び両方とも変異した株を用いて依存性を調べる実験が行われた（Dubaquié ら、1998）。野生型酵母の細胞から抽出した mRNA を［^{35}S］メチオニン存在下で in vitro 翻訳し、ラベルされたポリペプチド鎖をそれぞれの変異株より調製したミトコンドリアへ輸送させた。その結果、変異株のミトコンドリアを用いて輸送させた時には凝集するタンパク質が生じることがわかった。この凝集したタンパク質を 2 次元電気泳動で解析した結果、Cpn60、Cpn10 を共に必要とするタンパク質、Cpn10 には非依存だが Cpn60 は必要であるものという 2 種類があった。また、Cpn60 自身の巻き戻りには Cpn10 を必要とするという結果も示された。

真核生物のシャペロニンシステム

シャペロニン

高等植物においては、シャペロニンのホモログはミトコンドリア（Prasad と Stewart、1992、Tsugeki ら、1992）と葉緑体（Martel ら、1990）に存在する。ミトコンドリアのシャペロニン（Cpn60、しばしば Hsp60 とも表記される）は細菌のものと非常に相同性が高い。葉緑体のシャペロニン（Cpn60）は元々、RuBisCO 結合タンパク質として同定された（Hemmingsen ら、1988）。葉緑体にタンパク質を輸送させると、ATP 依存的に Cpn60 と結合することが Rieske 鉄—硫黄タンパク質（Madueño ら、1993）やフェレドキシン NADP$^+$還元酵素（Tsugeki と Nishimura、1993）などで示されている。ミトコンドリアのホモログが 1 種類のポリペプチド鎖で機能するのに対し、葉緑体のホモログは配列上明らかに異なる 2 種類のポリペプチド鎖（αとβ）がほぼ同量存在しており（Zabaleta ら、1992）、Cpn60αと Cpn60βはヘテロな多量体を形成していることが示されている（Nishio ら、1999）。2種類の Cpn60 の違いを調べるため、*Brassica napus* やエンドウの Cpn60αと Cpn60βは大腸菌で成熟型タンパク質として発現させて解析された（Clony ら、1992、Dickson ら、2000）。Cpn60βは Cpn60αに依存せず 14 量体を形成することができたのに対し、Cpn60αは Cpn60β存在下でのみヘテロな 14 量体を形成できることがわかった。このように Cpn60αと Cpn60βは単にアミノ酸配列が異なるだけでなく、機能的にも何らかの違いがあることが示されている。

コ-シャペロニン

ミトコンドリアのコ-シャペロニンホモログ（Cpn10）は大腸菌 GroES の機能を相補するものとして哺乳動物でまず見つかった（Lubben ら、1990)。本研究では、GroES の温度感受性突然変異株を用いて植物のミトコンドリアの Cpn10 をクローニングすることができた（Koumoto ら、1996)。

また、GroEL と共にタンパク質の巻き戻りを助ける活性を指標として、エンドウの葉緑体の抽出物からコ-シャペロニンホモログの存在が示された（Bertsch ら、1992)。エンドウのコ-シャペロニンホモログの部分アミノ酸配列を元に、ホウレンソウホモログの cDNA が単離された（Bertsch ら、1992)。その結果、葉緑体のコ-シャペロニンホモログはトランジットペプチドの後ろに 2 つの GroES（Cpn10）様ドメインがタンデムにつながった構造であることがわかった。成熟型のタンパク質の分子量は約 20 kDa と GroES の倍の大きさである（このため Cpn20 と表記される)。成熟型 Cpn20、アミノ末端側 GroES 様ドメイン及び、カルボキシル末端側 GroES 様ドメインは、大腸菌の GroES の変異株を相補できることが確かめられたが、1 ドメインのみでは全長に比べ機能が低いことが示された（Baneyx ら、1995、Bertsch と Soll、1995)。また、in vitro ではどちらのドメインも GroEL に対してコ-シャペロニンとしての活性を示すことはできなかった（Bertsch と Soll、1995)。本研究ではシロイヌナズナの Cpn20 の cDNA を単離し、成熟型として大腸菌で大量に発現させることができた。その発現産物を用いて解析した結果、Cpn20 は 4 量体で機能することが明らかとなった（Koumoto ら、1999)。

葉緑体のコ-シャペロニンホモログに関しては、Schlicher と Soll（1996）が葉緑

体のタンパク質について Cpn20 の部分ペプチドに対して作製された抗体を用いて行ったイムノブロット解析より新しい知見が得られた。ストロマ画分には Cpn20 に相当する 20 kDa のバンドが検出され、さらにチラコイド内腔画分には 10 kDa のバンドが検出された。この結果から、チラコイド内腔画分に Cpn10 が存在している可能性が示唆された。本研究では、シロイヌナズナの EST データベースをホモロジー検索することにより、葉緑体局在型の Cpn10 の cDNA を単離することができた（Koumoto ら、投稿中）。葉緑体の Cpn10 は既知のコ-シャペロニンホモログとは相同性が低く新奇なものであった。

以上のように本研究ではシロイヌナズナにおいて局在の異なる 3 つのコ-シャペロニンホモログを同定した。特に葉緑体に局在するホモログは他とは異なる特徴が見られたので、そのことについて考察した。

材料と方法

cDNA ライブラリーの構築

シロイヌナズナの cDNA ライブラリーは、大腸菌の発現ベクター pBluescript™（Stratagene 社）を用いて、Mori ら（1991）のベクター・プライマー法に準じて構築されたものを使用した。Poly(A)$^+$RNA は 7 日間暗所で育てたシロイヌナズナ（*Arabidopsis thaliana*, ecotype Landsberg erecta）から調製した。

ヒスチジンタグを持つタンパク質の発現と精製

各タンパク質の cDNA を発現ベクター（pQE30,32, QIAGEN 社）のいずれかにタンパク質の読み枠があうように挿入した。ヒスチジン 6 残基（ヒスチジン・タグ）は発現ベクターに由来しており、各発現タンパク質のアミノ末端に位置している。作製したプラスミドを用いて大腸菌を形質転換し、IPTG（isopropyl-β-D-thiogalactopyranoside）を用いてタンパク質の発現を誘導した。大腸菌は超音波破砕し、その抽出物を Ni-NTA（Ni-nitrilo-triacetic acid, QIAGEN 社）カラムもしくは Ni をカップリングさせた HiTrap キレーティングカラム（Pharmacia 社）にかけた。ヒスチジンタグを持つタンパク質はこれらのカラムに吸着するので、非吸着成分を十分洗い流した後、イミダゾールのグラジエント（0-0.5 M, 20-30 ml）をかけ溶出した。

植物の形質転換

バイナリーベクター pBI121Hm のβ–グルクロニダーゼ遺伝子を各タンパク質の

cDNA と置換し、プラスミドを作製した。pBI121Hm は pBI121 由来で、薬剤抵抗性遺伝子を２つ含んでいる（ネオマイシンフォスフォトランスフェラーゼ II 遺伝子とハイグロマイシンフォスフォトランスフェラーゼ遺伝子)。作製したプラスミドはアグロバクテリウム（*Agrobacterium tumefaciens*, strain EHA101）に導入し、植物を形質転換するのに用いた。タバコを形質転換する際にはリーフディスク法(Horsch ら,1985)を用いた。シロイヌナズナ（*Arabidopsis thaliana*, ecotype Columbia）は *in planta* 法を用いて、形質転換を行った（Bechtold ら,1993)。

電気泳動とイムノブロッティング

SDS-PAGE は Laemmli（1970）の方法に従った。ゲル上で分離されたタンパク質はセルロース膜または PVDF（polyvinylidene difluoride）膜に電気的に転写した。免疫化学的な検出には ECL detection system （Amersham 社）を用いた。検出後、PVDF 膜は水で洗浄してから CBB で染色した。抗体はそれぞれのタンパク質にヒスチジンタグをつけ大腸菌で発現させたものを抗原として作製した。60-80 kDa のタンパク質をよく分離するためには 0.15%のビスアクリルアミドを含む 15%ゲルを用いた。非変性の PAGE を行う場合は、Laemmli の方法から SDS のみを除いて行い、変性を防ぐため 4℃で泳動した。泳動後 BPB の線でゲルを切ってから CBB で染色した。タンパク質の移動度を BPB の移動度で標準化した相対的な移動度（Relative mobility, Rm）を測定し、Hedrick と Smith（1968）の方法通りゲル濃度に対してプロットした。キャリブレーション曲線を作成するためのマーカーとして、キモトリプシノーゲン（25 kDa)、BSA（67 kDa)、GroES（73 kDa)、カタラーゼ（232 kDa)、フェリチン（440 kDa）を使用した。

細胞分画

細胞分画は Tsugeki らの方法に準じて行った（1992）。試料としてカボチャの子葉又はタバコ、シロイヌナズナの本葉を使用した。試料 1 g に対し抽出バッファー（150 mM Tricine-KOH, pH7.5, 13% (w/w) sucrose, 1 mM EDTA）を 3 ml 加え、かみそりの刃で細かくチョッピングした。この破砕液を 3 枚に重ねたガーゼでこし、1 mM EDTA を含む 30-60% (w/w)のショ糖密度勾配の上に重層した。遠心は 85500 × g で 2.5 時間 4℃で行った。遠心後、遠沈管の底から試料を 0.5 ml ずつの画分に分けて回収した。ミトコンドリアのマーカー酵素としてチトクローム c 酸化酵素（Hodges と Leonard、1974）を、葉緑体のマーカーとしてトリオースリン酸イソメラーゼ（Feierabend、1975）を、マイクロボディのマーカーとしてカタラーゼ（Luck、1965）を用い、それぞれの報告されている方法に従い活性を測定した。クロロフィルの含有量は Arnon の方法（1949）に従い測定した。

免疫電顕法

Nishimura ら（1993）の方法に従って、固定と脱水を行った葉を LRWhite レジン（London Resin Co.社）に包埋した。免疫染色の方法は Kinoshita ら(1999)の方法に準じて行った。用いた一次抗体濃度は、AT1 に対する抗体が 500 倍希釈、AT2 に対する抗体は 200 倍希釈、AT3 に対する抗体は 50 倍希釈、RuBisCO に対する抗体は 5000 倍希釈である。Protein A-gold（Amersham 社）はすべて 30 倍希釈で用いた。染色した試料は透過型電子顕微鏡（1200, JEOL）を用い、80 kV

で観察した。

ノザン解析

全 RNA は ISOGEN（ニッポンジーン社）で抽出し、塩化リチウム沈殿により精製した。全 RNA10 μg 分（AT3 を検出する場合は 20 μg 分）をホルムアルデヒドを含む 1%アガロースゲルで分離し、真空転写装置を用いてナイロン膜（HibondN$^+$、Amersham 社）に転写した。転写には 1 N の酢酸アンモニウムを用いた。転写後、UV 処理により RNA を膜に固定した。プローブを作製するために各プラスミドの cDNA 部分を PCR で増幅し、増幅された DNA 断片を BcaBEST labelling kit（宝酒造社）を用いて[α-^{32}P]dCTP で放射線標識した。ハイブリダイゼーションは Church と Gilbert（1984）の方法に従った。ハイブリダイゼーション後、まず膜を 0.1% SDS を含む 2 倍濃度の SSC 液（150 mM NaCl, 15 mM sodium citrate, pH7.0）中で 42℃30 分間洗浄し、続けて 0.1% SDS を含む 0.1 倍濃度の SSC 液中で 65℃（AT3 のみ 55℃）で 1 時間以上洗浄した。風乾後膜をイメージングプレートにコンタクトし、放射線のシグナルを Bio-Imaging Analyzer System（Fuji 社）で解析した。

ゲルろ過カラムクロマトグラフィー

Superose12HR 10/30 カラムを FPLC（Pharmacia 社）システムで使用した。カラムはバッファー A（100 mM Tris-HCl, pH 7.7, 10 mM $MgCl_2$, 10 mM KCl）で平衡化した。流速は 0.5 ml/min で行い、280 nm の吸光度をモニターした。カタラーゼ（232 kDa）、BSA（67 kDa）、リボヌクレアーゼ A（13.7 kDa）、ビタミン B_{12}

（1355 Da）を分子量のマーカーとして用いた。HiLoad 16/60 Superdex 200pg カラムは AKTA（Pharmacia 社）システムで使用した。50 mM の NaCl を含むバッファー A を用い、流速 0.8 ml/min で行った。分子量マーカーとして用いたのは、上記の 4 種類に加え、チオグロビン（669 kDa）、フェリチン（440 kDa）、アルドラーゼ（158 kDa）、オバルブミン（43 kDa）、キモトリプシノーゲン（25 kDa）の計 9 種類である。溶出パラメータの Kav と分子量の対数を用い、キャリブレーション曲線を作成した。Kav 値はそれぞれのタンパク質に対し、公式 Kav = (Ve-Vo)/(Vt-Vo)に従い計算した。この時 Ve はそれぞれのタンパク質の溶出に要した液量、Vo はカラムのボイドボリューム、Vt はカラムのベッドボリュームを示している。

DSP（Dithiobis-succinimidyl propionate）を用いた架橋実験

オリゴマー構造を解析するため、DSP により架橋されたオリゴマーを調製した。DSP は 50 mg/ml になるように DMSO に溶かした。この DSP 溶液を 1 ml のタンパク質溶液に対し 5 µl 加えた。タンパク質溶液はヒスチジンタグを付加した AT2 タンパク質が約 100 µg と 0.1 M NaCl を含む 50 mM リン酸バッファー（pH 7.2）である。この混合液を 4℃で一晩反応させた。5 µl の 10 M 酢酸アンモニウムを加え架橋を停止させた。タンパク質は 10%(w/v)の TCA（trichloroacetic acid）で沈殿させ、SDS-PAGE 用のサンプルバッファーで溶解し、SDS-PAGE に用いた。 DSP は分子内にジスルフィド結合が 1 つあるため、2ME（2-mercaptoethanol）存在下ではモノマーが、非存在下では架橋されたオリゴマーが検出される。

質量分析

ポジティブイオン MALDI-TOF 質量分析は Voyager Elite XL time-of-flight 質量分析計（Perseptive Biosystems 社）で Asahi ら（1997）の方法に従って行った。架橋された AT2 オリゴマーは逆相カラムクロマトグラフィーにより精製し、試料とした。試料はステンレスのプレート上に乗せ、マトリックス溶液（シナピン酸で飽和した 33%アセトニトリルの上清）と混合し、風乾してから測定した。イオンは窒素レーザー（337 nm）により生成した。

シャペロニンを介したタンパク質巻き戻り実験

シャペロニンを介したタンパク質巻き戻り実験は Schmidt ら（1994）の方法に準じて行った。クエン酸合成酵素（CS, 15 μM）を 6 M 塩酸グアニジンと 20 mM DTT を含む 100 mM のトリスバッファー （pH 7.7）で溶解し、室温で 1 時間以上放置することにより変性させた。変性した CS は 再生バッファー（10 mM $MgCl_2$, 10 mM KCl, 225 nM GroEL 14 量体, 50 mM Tris-HCl, pH7.7）で 150 nM になるように素早く希釈し、0℃においた。温度を 35℃にすると同時に 2 mM ATP と各種コシャペロニンをそれぞれ加えた。種々の反応時間後一部を取り出し、CS の活性を Srere（1969）の方法に従い測定した。GroEL と GroES は宝酒造社から CS はベーリンガー社から購入した。コシャペロニンのタンパク質量は Protein assay kit （BioRad Laboratories 社）を用いて決定した。

GroEL-AT2 複合体の解析

2.1 μM の GroEL14 量体（$GroEL_{14}$）と 1.3 μM のヒスチジンタグを付加した AT2

の 4 量体（$AT2_4$）をバッファー B（100 mM Tris-Cl,pH7.7, 10 mM $MgCl_2$, 10mM KCl, 50 mM NaCl, 1 mM ADP）中で室温で 1.5 時間反応させることにより GroEL-AT2 複合体を生成させた。GroEL-AT2 複合体の精製は複合体が安定に存在する 1 mM ADP を含んだ条件下で行った。まず、複合体になっていない $AT2_4$ をゲルろ過カラムを用いて除いた。次に複合体になっていない $GroEL_{14}$ を除くため、Ni をカップリングさせた HiTrap キレーティングカラム（Pharmacia 社）を用いた。複合体になっていない $GroEL_{14}$ はヒスチジンタグを持たないため保持されないが、GroEL-AT2 複合体はヒスチジンタグを付加した AT2 を含むので吸着する。複合体は 500 mM のイミダゾールで溶出した後、逆相カラムクロマトグラフィー（Sephacil Protein C4 5mm ST 4.6/250, Pharmacia 社）により GroEL と AT2 の画分としてそれぞれ回収し、アミノ酸分析を行った。アミノ酸分析は 0.2%フェノールを含む 6 N 塩酸で気相加水分解した後、L-8500 A amino acid analyzer（日立社）により加水分解物のアミノ酸含量を定量した。

<u>定量的 PCR</u>

鋳型になる cDNA は Superscript™ preamplification system for first strand cDNA synthesis kit（Gibco BRL 社）を用い、全 RNA より作製した。逆転写はオリゴ（dT）プライマーを用いて行った。定量的 PCR は ABI Prism 7700 sequence detection system（PE Biosystems 社）を使用して行った。この機器は以下のような原理を利用することにより、定量的 PCR を行う。特異的な検出のために PCR で増幅される領域に相補的に結合する蛍光プローブを作製しておく。この蛍光プローブには 5'末端にフルオレセイン系の蛍光色素（リポーター）を、3'末端

にはローダミン系の蛍光色素（クェンチャー）を付加しておく。プローブはそのままの状態では蛍光共鳴エネルギーがリポーターからクェンチャーに移動するのでリポーターの蛍光は抑制されている。標的部分に相補的に結合したプローブは、PCR の伸長反応の際 DNA ポリメラーゼにより分解される。この分解によりリポーターはクェンチャーと遊離することになり、リポーターの蛍光強度が増加する。リポーター由来の蛍光強度を PCR の各過程を通じてモニターし、PCR 産物の増幅プロットを作製することにより、PCR 産物の量を定量化することができる。PCR プライマーと蛍光プローブを作製する際にはコンピュータープログラム Prime Express™ (PE Biosystems 社) を用いた。PCR 反応は変性は 95℃で 15 秒、アニーリングと伸長は 60℃で 1 分という設定で 40 サイクル繰り返すプログラムで行った。

結果

シロイヌナズナのコーシャペロニンホモログ

大腸菌 GroES 欠損株を用いたコーシャペロニンホモログの単離

タンパク質は正しい高次構造をとることによって、初めてその機能を果たすことができる。高次構造形成を形成する時の最大の問題は、タンパク質の分子内で結合するべきドメインが分子間で結合してしまうために凝集し、非可逆的に変性してしまうことである。この凝集を防ぐため、生物はシャペロニンシステムという 1 分子ごとの高次構造形成の‘場’を持っている。植物のシャペロニンシステムについて調べるため、まず、その制御因子であるコ-シャペロニンの解析を行った。シロイヌナズナのコーシャペロニンホモログの cDNA クローンを単離するため、Functional cloning を行った。大腸菌の YJ004（KI471）株はコ-シャペロニンをコードする GroES に欠損を持ち、30℃では生育できるが 42℃では生育できない温度感受性突然変異株である。この株を発現ベクター pBluescript ™ で作製したシロイヌナズナの cDNA ライブラリーを用いて形質転換し、42℃で生育できるようになった大腸菌をスクリーニングした。1 次スクリーニングで陽性のクローンを用いて再度形質転換を行い、2 次スクリーニングを行った。図 3 で示すように、変異株（2）やネガティブコントロール（ベクターのみ、3 と 4）は 42℃では生育できなかったが、2 個の陽性クローン（5 と 6）は野生型（1）と同様 42℃でも生育可能であった。陽性クローンから単離したプラスミドに挿入されていた cDNA 断片の塩基配列を決定しホモロジー

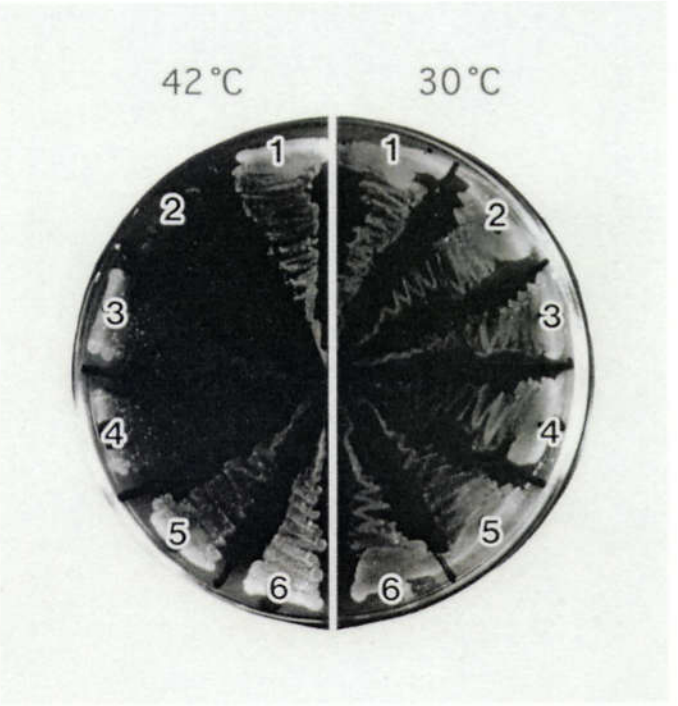


図3　大腸菌のGroES欠損温度感受性突然変異株を相補するコーシャペロニンホモログのcDNAクローンの単離
4種類の大腸菌を2枚の寒天培地にストリークした。1枚は30℃（右）で、もう1枚は42℃（左）で培養した。1は野生型、2はGroES欠損温度感受性変異体YJ004（KI471）株、3と4はベクターpBSのみを導入した変異体、5と6はコーシャペロニンホモログのcDNAを持つ変異体を示す。

検索を行ったところ、全て 10 kDa のコーシャペロニン（Cpn10）のホモログをコードしていることがわかった。

Cpn10 ホモログ（AT1）の塩基配列とアミノ酸配列

Cpn10 ホモログ（AT1 と呼ぶ）の塩基配列と推定されるアミノ酸配列を図 4a に示した。cDNA は全長 647 塩基対で、98 アミノ酸のタンパク質をコードしてい た。3'末端の非翻訳領域は 340 塩基であり、Grellet ら（1993）の報告より長かったが、その他の領域では一致していた。推定アミノ酸配列のアミノ末端 17 残基は両親媒性のαヘリックスをとることが予想され、ミトコンドリアへの輸送シグナルとして機能することが示唆された（図 4b）。ただし、このシグナルは延長配列ではなく、切断は受けないと考えられた。また、大腸菌のコ-シャペロニン（GroES）とは同一性が低い（30%）のに対し、酵母や動物のミトコンドリア Cpn10 により高い同一性（それぞれ 39%と 50%）を示したことからもミトコンドリア Cpn10 のホモログであると予想された。

Cpn10（AT1）の局在性

Cpn10（AT1）の局在性を確かめるため、シロイヌナズナのロゼット葉を用い細胞分画を行った。ショ糖密度勾配遠心により細胞分画を行い、マーカー酵素の活性を測定してオルガネラの分布を確認した。イムノブロット解析を行った結果、Cpn10（AT1）はミトコンドリアのマーカー酵素と同じ分布を示した（図 5A）。さらに、Cpn10（AT1）の局在を免疫電顕により確かめた。シロイヌナズナの内在性の Cpn10（AT1）は検出することができなかったので、Cpn10（AT1）を過

(A)

```
                                    gaagaattgagaa   13
ATGATGAAGCGTCTGATCCCAACGTTCAAC      43
 M  M  K  R  L  I  P  T  F  N       10
CGCATCTTGGTGCAGAGAGTCATCCAGCCC      73
 R  I  L  V  Q  R  V  I  Q  P       20
GCTAAAACCGAAAGCGGCATTCTCCTACCT     103
 A  K  T  E  S  G  I  L  L  P       30
GAGAAATCCTCCAAGCTGAACTCAGGCAAG     133
 E  K  S  S  K  L  N  S  G  K       40
GTGATAGCTGTTGGACCTGGATCAAGGGAT     163
 V  I  A  V  G  P  G  S  R  D       50
AAGGACGGGAAATTGATTCCGGTCTCTGTG     193
 K  D  G  K  L  I  P  V  S  V       60
AAGGAAGGCGACACTGTTCTTCTTCCAGAG     223
 K  E  G  D  T  V  L  L  P  E       70
TACGGTGGTACACAGGTCAAGCTCGGCGAG     253
 Y  G  G  T  Q  V  K  L  G  E       80
AACGAGTACCATCTCTTCCGGGACGAGGAT     283
 N  E  Y  H  L  F  R  D  E  D       90
GTTTTGGGAACTTTGCACGAGGATtgaaaa     313
 V  L  G  T  L  H  E  D  *          98
ggctaagcttgccaacttaaccacgagggt     343
tcatgttggtgtttgtgttatgaggagaag     373
tcatttataaattagtttatcttgaagatg     403
tggttggactttgttgtcgtttatcattga     433
atctacctttatgaacctgtctttgaattt     463
ttacaaatgggcatcaatcacatggataac     493
ccaagtgttgcatcttctcattttgtgctt     523
ttccgtaatctgtggatgcttttcgtttac     553
gtttaggcagcaacttcaatagccaatgta     583
gacgatagcttagttgactttgcttattca     613
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa     643
aaaa                               647
```

(B)

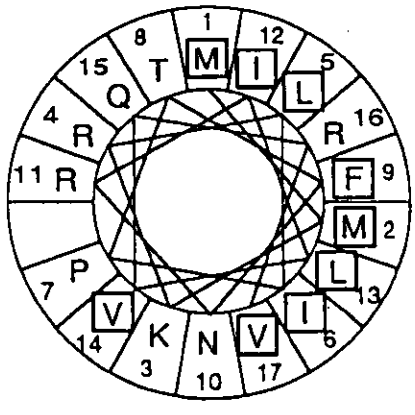


図 4 Cpn10 ホモログ（AT1）の構造（A）とアミノ末端 17 残基の予測されるαヘリックス構造（B）

（A） Cpn10 ホモログ（AT1）の塩基配列と推定されるアミノ酸 1 次構造を示した。塩基は 5' 末端から番号をつけた。推定されるアミノ酸の配列は最初のメチオニン残基から数えて番号をつけた。★印は停止コドンを示す。

（B） 予測されるαヘリックス構造（1 周 3.6 残基）を軸方向から示した。疎水性残基は四角で囲んである。

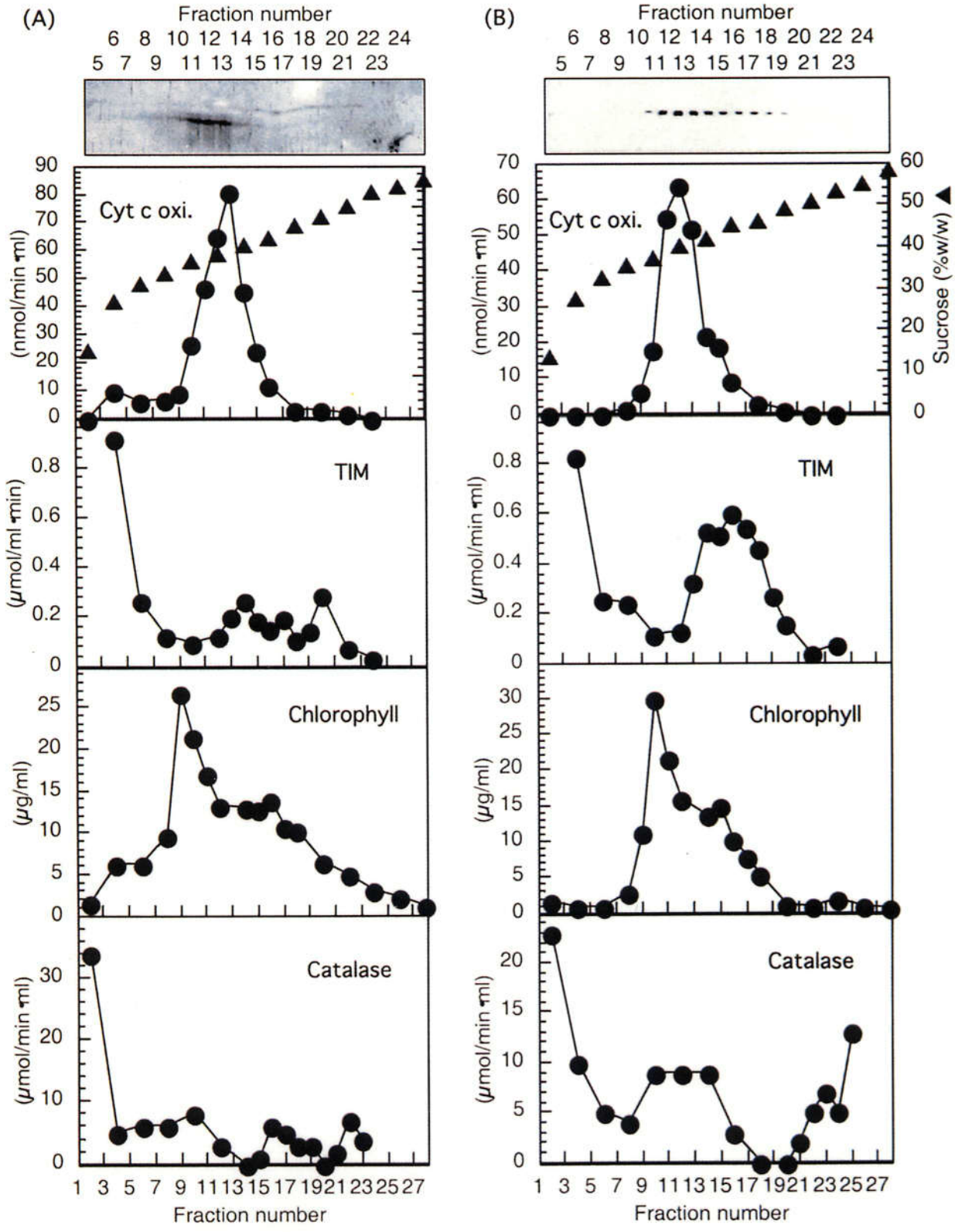


図5　ショ糖密度勾配遠心による細胞分画

（A）シロイヌナズナのロゼット葉を用いて細胞分画を行った。各分画をSDS-PAGEにかけCpn10（AT1）に対する抗体を用いてイムノブロット解析を行った（上）。ミトコンドリアのマーカー酵素としてチトクロームc酸化酵素（Cyt. c oxi.）、葉緑体のマーカー酵素としてトリオースリン酸イソメラーゼ（TIM）、チラコイド膜のマーカーとしてクロロフィル（Chlorophyll）、マイクロボディのマーカー酵素としてカタラーゼ（Catalase）の活性又は含有量を測定した（下）。

（B）Cpn10（AT1）を過剰発現する形質転換タバコの葉を用い（A）と同様に細胞分画を行った。

剰発現する形質転換タバコを作製した。Cpn10（AT1）の cDNA 全長をカリフラワーモザイクウイルスの 35S プロモーターの下流につなぎ、タバコに導入した。Cpn10（AT1）の発現による植物への影響は特に認められなかった。Cpn10（AT1）に対する抗体はタバコの内在性の Cpn10 ホモログにはほとんど反応しなかったので、イムノブロット解析においては導入した Cpn10（AT1）のみが検出された（図 6a)。

Cpn10 ホモログ（AT1）を過剰に発現しているライン（図 6a の＃2）を用いて細胞分画および免疫電顕を行った。細胞分画の結果から、導入した Cpn10（AT1）は形質転換タバコにおいてもミトコンドリアのマーカー酵素と同じ分布を示し、正しく輸送されていることがわかった(図 5B)。また、図 6b に示すように、免疫電顕像においてもミトコンドリアに金粒子が局在することが示された。ベクターのみを導入したタバコ植物においては金粒子は検出されないことから、金粒子は導入した Cpn10（AT1）の局在を示していることが確かめられた。これらの結果より、Cpn10（AT1）はミトコンドリアに局在することが明らかとなった。

<u>葉緑体局在型コ-シャペロニンホモログ、Cpn20</u>

植物のミトコンドリアのシャペロニンシステムは、動物のミトコンドリアのものや真正細菌のものと非常に類似していることがわかった。植物細胞には共生によって生じたと考えられるもう 1 つのオルガネラ、葉緑体が存在する。葉緑体のコ-シャペロニンホモログである Cpn20 は、ホウレンソウにおいて初めて cDNA クローンが単離された（Bertsch ら、1992)。Cpn20 は Cpn10（GroES）様

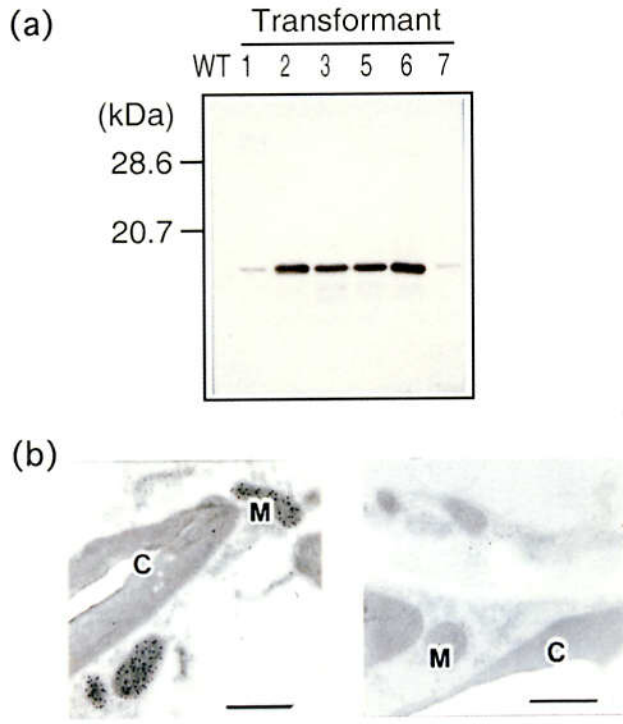


図6　Cpn10（AT1）を過剰発現する形質転換タバコのイムノブロット解析（a）と免疫電顕（b）

（a）　形質転換タバコと野生型タバコの葉の抽出物を遠心し、上清のタンパク質量を測定した。それぞれ10 μgのタンパク質をSDS-PAGE にかけ、His- Cpn10（AT1）に対する抗体を用いてイムノブロット解析を行った。分子量マーカーの位置を左に示した。

（b）　Cpn10（AT1）を過剰発現する形質転換タバコ（左）とベクターのみを導入したタバコ（右）は、花芽から数えて6枚目の葉を用い免疫電顕を行った。バーの長さは0.5 μmで、Mはミトコンドリア、Cは葉緑体を示している。

のドメインが 2 つタンデムにつながった構造をもつ。ホウレンソウのクローンの塩基配列を元にシロイヌナズナの EST データバンクを検索したところ、カルボキシル末端に高い相同性を示す EST クローン（accession number EMBL Z18060）が見つかった。そこで、この EST クローンの塩基配列に基づいてプライマーを作製し、PCR で増幅した断片をプローブとして、シロイヌナズナの cDNA ライブラリーより Cpn20 の cDNA クローンの単離を試みた。単離した cDNA は 958 塩基対の長さで、253 アミノ酸をコードしていた（図 7、AT2 と呼ぶ)。ホウレンソウの Cpn20 との比較より、アミノ末端の 51 アミノ酸が葉緑体 への移行シグナルとして働くトランジットペプチドであると予想された。予想される成熟タンパク質部分のホウレンソウとシロイヌナズナとの同一性は 61%であった。シロイヌナズナのミトコンドリアの Cpn10 との比較では、アミノ末端の Cpn10（GroES）様のドメイン（60-154 残基）が 30%、カルボキシル末端のドメイン（158-253 残基）が 37%の同一性を示した。2 つのドメインの同一性は 42%であった。

Cpn20（AT2）の局在性

Cpn20（AT2）の局在を免疫電顕により確かめた。Cpn10（AT1）の時と同様、Cpn20（AT2）を過剰発現する形質転換タバコを作製した。図 8a に Cpn20（AT2）を導入した形質転換体 6 ラインとベクターのみを導入したもののイムノブロット解析の結果を示した。ベクターのみを導入した植物体において、Cpn20（AT2）に対する抗体に反応する、非常に近接した 2 本のバンドが検出された。このことはタバコには Cpn20 が少なくとも 2 種存在することを示唆していると考えら

```
aatttttagggttttatcctccgaaagtctcaacctttttcttatcctcaacaaggagaa     60
ATGGCGGCGACTCAACTTACAGCGTCACCAGTGACTATGTCAGCAAGGAGCTTAGCCTCG    120
M  A  A  T  Q  L  T  A  S  P  V  T  M  S  A  R  S  L  A  S      20
CTGGATGGTCTCAGAGCTTCGAGTGTCAAGTTTTCATCTTTGAAACCAGGGACCCTTAGA    180
L  D  G  L  R  A  S  S  V  K  F  S  S  L  K  P  G  T  L  R      40
CAGAGCCAGTTCCGTCGTTTGGTTGTCAAAGCTGCTTCTGTTGTTGCCCCTAAGTATACT    240
Q  S  Q  F  R  R  L  V  V  K  A  A  S  V  V  A  P  K  Y  T      60
TCAATTAAGCCATTGGGAGATCGAGTTTTGGTGAAGATCAAGGAGGCAGAGGAGAAGACT    300
 S  I  K  P  L  G  D  R  V  L  V  K  I  K  E  A  E  E  K  T      80
TTAGGTGGTATCTTACTTCCATCCACTGCTCAATCAAAACCTCAAGGAGGTGAAGTCGTT    360
 L  G  G  I  L  L  P  S  T  A  Q  S  K  P  Q  G  G  E  V  V     100
GCCGTGGGTGAAGGAAGAACTATTGGGAAGAACAAAATTGATATCACTGTCCCTACTGGA    420
 A  V  G  E  G  R  T  I  G  K  N  K  I  D  I  T  V  P  T  G     120
GCACAAATTATCTACTCCAAATACGCAGGAACTGAGGTGGAGTTCAATGATGTGAAGCAT    480
 A  Q  I  I  Y  S  K  Y  A  G  T  E  V  E  F  N  D  V  K  H     140
CTTATCCTCAAGGAAGATGATATTGTTGGCATTCTTGAGACAGAGGACATCAAAGATCTC    540
 L  I  L  K  E  D  D  I  V  G  I  L  E  T  E  D  I  K  D  L     160
AAACCTTTGAATGACCGAGTCTTTATTAAGGTTGCTGAGGCGGAGGAGAAAACAGCTGGA    600
 K  P  L  N  D  R  V  F  I  K  V  A  E  E  A  E  E  K  T  A  G     180
GGGTTGTTGTTAACCGAGACTACCAAAGAGAAGCCTTCTATTGGCACGGTGATAGCAGTT    660
 G  L  L  L  T  E  T  T  K  E  K  P  S  I  G  T  V  I  A  V     200
GGACCGGGTTCCCTAGACGAGGAAGGTAAAATTACGCCTCTACCAGTATCAACCGGAAGC    720
 G  P  G  S  L  D  E  E  G  K  I  T  P  L  P  V  S  T  G  S     220
ACAGTACTTTACTCCAAGTATGCTGGTAACGACTTCAAGGGCAAAGATGGTTCCAACTAC    780
 T  V  L  Y  S  K  Y  A  G  N  D  F  K  G  K  D  G  S  N  Y     240
ATTGCCCTCAGAGCTTCAGATGTGATGGCTATACTTTCTTAGttatgttatatctttgta    840
 I  A  L  R  A  S  D  V  M  A  I  L  S  *                       253
atctgcaacttgtatcccaattgtggaaatttttttccgtaaacggcctgaacataatctg    900
gaataaagacttgagtttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa    958
```

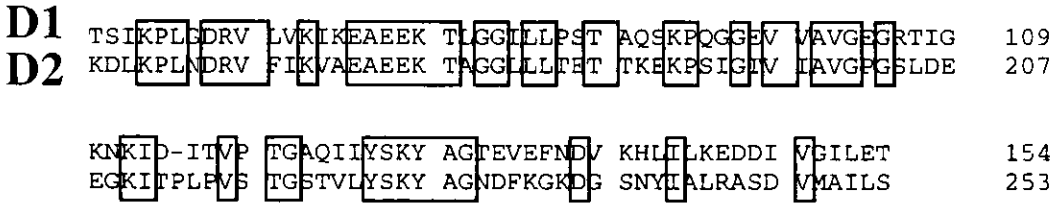


図7　Cpn20（AT2）の構造

(a) Cpn20（AT2）の塩基配列と推定されるアミノ酸１次構造を示した。塩基は5'末端から番号をつけた。推定されるアミノ酸の配列は最初のメチオニン残基から数えて番号をつけた。ホウレンソウのCpn20の比較から予想されるトランジットペプチド部分を白抜きで示した。★印は停止コドンを示す。

(b) アミノ末端のCpn10様ドメイン（残基60-154）とカルボキシル末端のCpn10様ドメイン（残基158-253）を比較した。同一のアミノ酸を四角で囲んで示した。

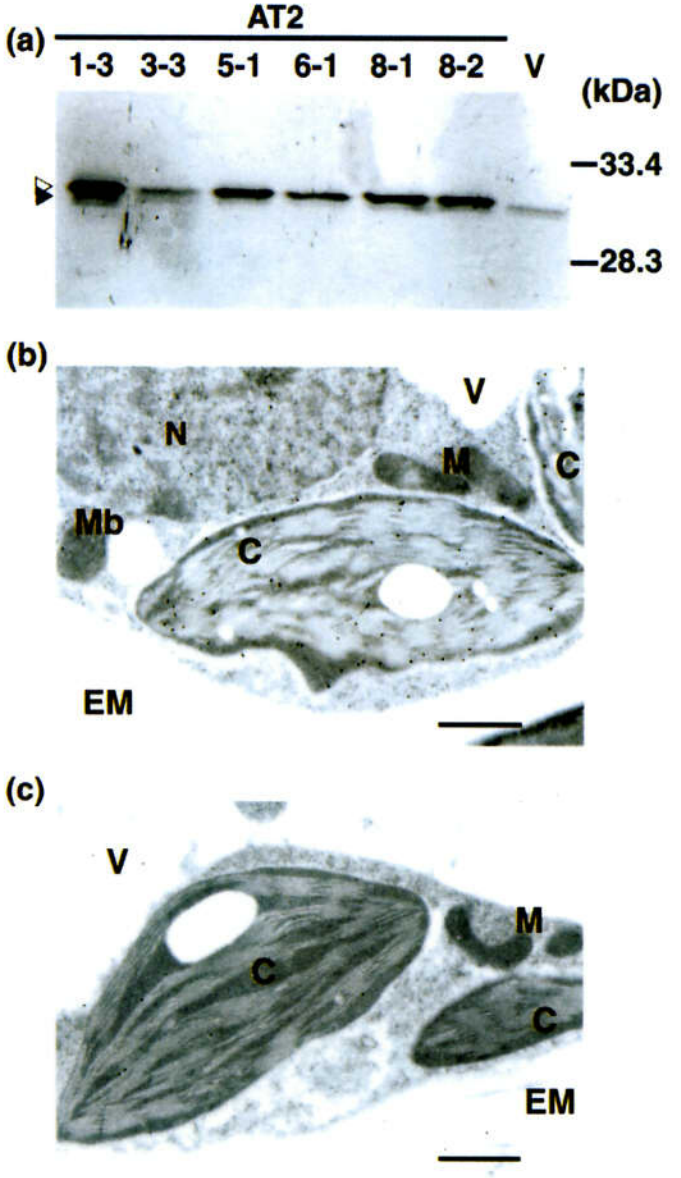


図8　Cpn20(AT2)を過剰発現する形質転換タバコのイムノブロット解析(a)と免疫電顕(b, c)

(a)　形質転換タバコの葉の抽出物を遠心し、上清のタンパク質を測定した。それぞれ10μgのタンパク質をSDS-PAGEにかけ、His-Cpn20(AT2)に対する抗体を用いてイムノブロット解析を行った。番号は形質転換タバコのラインを示しており、Vはベクターのみを導入した植物体を示す。分子量マーカーの位置を左に示す。内在性のCpn20の位置を▲で、導入したシロイヌナズナのCpn20を△で示した。

(b, c)　HIs-Cpn20(AT2)に対する抗体を用いた免疫電顕には、寒天培地で育てた形質転換植物の葉をそれぞれ1枚ずつ使用した。Cpn20(AT2)を過剰発現する形質転換タバコの免疫電顕像を(b)に、ベクターのみを導入したタバコを(c)に示した。バーの長さは1μmで、Mはミトコンドリア、Cは葉緑体、Mbはマイクロボディ、Nは核、Vは液胞、EMは細胞外マトリックスを示している。

れる。導入したシロイヌナズナの Cpn20 はタバコの内在性のものとわずかに移動度が異なっていた。発現量はラインにより差が見られたが、Cpn20（AT2）の発現による植物への影響は特に認められなかった。

Cpn20（AT2）を過剰発現する形質転換タバコの免疫電顕像を図 8b に示した。金粒子は葉緑体にのみ見られ、他のオルガネラには見られなかった。ベクターのみを導入した植物体では金粒子はどこにも検出されなかったので（図 8c)、図 8b で見られる金粒子は導入したシロイヌナズナの Cpn20（AT2）の局在を示していることがわかった。このことから、Cpn10（AT1）と同様、Cpn20（AT2）もタバコ植物体の中で葉緑体へ正しく輸送されていることが確かめられた。

葉緑体局在型 Cpn10

上述のように、シロイヌナズナにおいて、ミトコンドリア Cpn10（AT1）と葉緑体 Cpn20（AT2）の塩基配列を決定することができた。2 つのコ-シャペロニンの成熟タンパク質部分のアミノ酸配列を用いて EST データバンクを検索したところ、どちらにも低い相同性を示すクローンが存在していた。EST クローン（GenBank accession number T44192）の塩基配列を決定したところ、開始メチオニンを含んでいた。cDNA は全長 603 塩基対で 139 アミノ酸から成るタンパク質をコードしていた（図 9a)。分子量並びにコーシャペロニンのアミノ末端に見られる保存モチーフ、PXX(D/N)(K/R)から判断すると、このクローンはアミノ末端に延長配列を持つ Cpn10 ホモログであると考えられた。この Cpn10 ホモログ（AT3 と呼ぶ）と他のコーシャペロニンとを比較するため、系統樹（図 10）を作成した。Cpn10（AT3）はアミノ末端の 49 残基を削り、アミノ末端の

(a)

```
          cccacgcgtccgggaagaagaagatctgcttctgcaaaaaaccttatcctgct      53
ATGGCTTCCACTTTCGTCTGCTCTCTACCAAATCCTTTCTTTGCTTTTCCGGTCAAAGCA     113
 M  A  S  T  F  V  C  S  L  P  N  P  F  F  A  F  P  V  K  A       20
ACTACTCCTTCGACGGCTAACTATACGCTTCTCGGAAGTCGAAGAGGTTGTCTTAGAATC     173
 T  T  P  S  T  A  N  Y  T  L  L  G  S  R  R  G  C  L  R  I       40
AAAGCGATTTCCACTAAATGGGAACCGACAAAGGTTGTTCCTCAGGCAGACAGAGTTCTT     233
 K  A  I  S  T  K  W  E  P  T  K  V  V  P  Q  A  D  R  V  L       60
GTTCGTCTTGAAGATCTTCCTATTAAATCCTCAGGTGGAGTATTGTTGCCTAAAGCAGCT     293
 V  R  L  E  D  L  P  I  K  S  S  G  G  V  L  L  P  K  A  A       80
GTGAAGTTTGAGAGATACCTAACAGGAGAGATTATATCTGTTGGTTCTGAGGTTGGACAA     353
 V  K  F  E  R  Y  L  T  G  E  I  I  S  V  G  S  E  V  G  Q      100
CAAGTTGGACCTGGAAAGAGGGTTTTGTTCTCTGATGTGAGCGCTTATGAGGTCGATTTG     413
 Q  V  G  P  G  K  R  V  L  F  S  D  V  S  A  Y  E  V  D  L      120
GGAACCGATGCTAGGCATTGCTTCTGTAAAGAGAGTGACTTGTTGGCCCTCGTTGAGTGA     473
 G  T  D  A  R  H  C  F  C  K  E  S  D  L  L  A  L  V  E  *      140
agtcttgtccaagagggagagatttgaagattttacaagttttctgtaattttcagacag     533
caattgttgtttctagttaatccttcaatttaatatcaattgagatcactttcagaaaaa     593
aaaaaaaaaa                                                       603
```

(b)

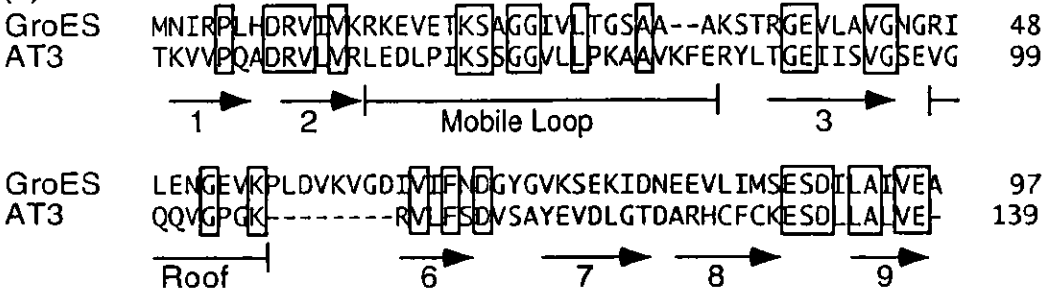


図9　Cpn10（AT3）の構造

(a) Cpn10（AT3）の塩基配列と推定されるアミノ酸１次を示した。塩基は5'末端から番号をつけた。推定されるアミノ酸の配列は最初のメチオニン残基から数えて番号をつけた。★印は停止コドンを示す。アミノ末端に見られる保存モチーフ、 PXX(D/N)(K/R)は四角で囲んで示した。

(b) Cpn10（AT3）の推定される成熟タンパク質部分と大腸菌のコ -シャペロニンであるGroESを比較した。結晶構造解析より観察された GroESの構造上の特徴をHuntらの報告（1996）から引用し、アミノ酸配列の下に示した。矢印はβストランドを示す。同一のアミノ酸は四角で囲んで示した。

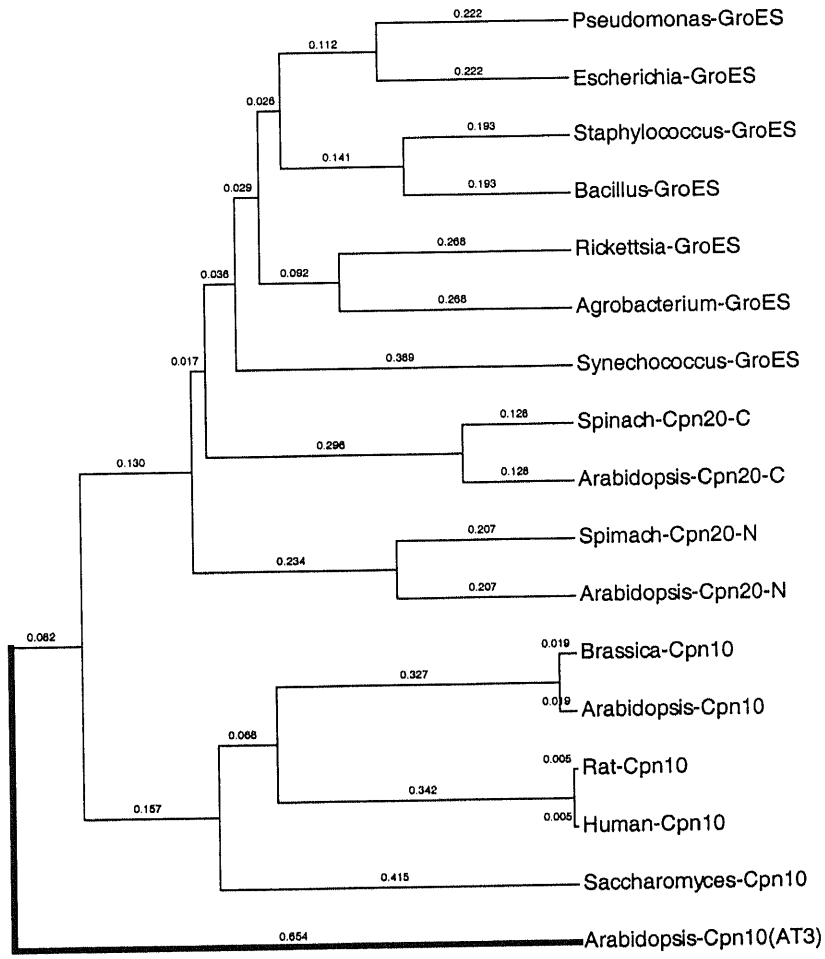


図10　コーシャペロニンホモログの系統樹

DNA解析プログラムGeneWorks内のUPGMAプログラムにより系統樹を作成した。それぞれのデータベース登録番号は、Pseudomonas-GroES, P30720; Escherichia-GroES, X07850; Staphylococcus-GroES, Q08841; Bacillus-GroES, P28599; Rickettsia-GroES, P80469; Agrobacterium-GroES, X68263; Synechococcus-GroES, M58751; Spinach-Cpn20, M87646; Brassica-Cpn10, U65890; Arabidopsis-Cpn10, D88314; Rat-Cpn10,P26772; Human-Cpn10, X75821; Saccharomyces-Cpn10, P38910である。Cpn10(AT3)とCpn20については以下に示す領域を用いた。Cpn10 (AT3), 残基50-139; Spinach-Cpn20-N, 62-157; Spinach-Cpn20-C, 160-255; Arabidopsis-Cpn20-N, 60-154; Arabidopsis - Cpn20-C, 158-253.

保存モチーフの位置を代表的なコ-シャペロニンである大腸菌の GroES と一致させた。図 10 に示した通り、Cpn10（AT3）は既知のコ-シャペロニンの中には高い相同性を示すものはなく、非常に新奇なものであることがわかった。

Cpn10（AT3）の局在性

Cpn10（AT3）の局在性を調べるため、まず PSORT プログラムを用いコンピュータによる予測を行った。その結果、アミノ末端側約半分（1-82 残基）が葉緑体ストロマへのトランジットペプチドであることが予測された。しかし、Cpn10（AT3）の予想されるトランジットペプチドの長さは約 40～50 アミノ酸であるため、いくつかの条件で再度予測を行った。その結果、カルボキシル末端を削除しトランジットペプチドと成熟タンパク質のアミノ末端側一部のみを含む配列を用いた場合、アミノ末端 26 残基がチラコイド内控への移行シグナルとして働くことが予測された。

次に、実際にシロイヌナズナのロゼット葉を用い細胞分画を行った。ショ糖密度勾配遠心にかけ各オルガネラを分離し、イムノブロット解析を行った。内在性の Cpn10（AT3）は量が少ないことと分解されやすい性質のため、バンドとして検出することができなかった。そこで、Cpn10（AT3）を過剰発現する形質転換シロイヌナズナを作製し、同様の実験を行った。過剰発現植物を用いても細胞分画後の各画分は濃度が低くなってしまい、Cpn10（AT3）のバンドを確認するためには多量のサンプルを SDS-PAGE にかける必要があった。シグナルは非常に弱いながら、Cpn10（AT3）のバンドは葉緑体の画分と葉緑体が壊れて生じたチラコイド膜を含む画分に見られた（結果省略）。このことから、Cpn10

（AT3）が葉緑体に存在することは確かめられたが、そのオルガネラ内の局在について明らかにすることはこの手法では難しいと考えられた。

そこで、Cpn10（AT3）を過剰発現する形質転換シロイヌナズナを用いて免疫電顕を行った。成熟した葉では葉緑体が完全に発達しており、チラコイド膜がほぼ葉緑体全体に広がりデンプンの蓄積も見られる。そのような葉緑体ではストロマの領域は極めて小さく、オルガネラ内の局在性を調べることは難しいと思われたので、若い子葉を用いて行った。シロイヌナズナの芽生えを 4 日間暗所で育てた後、連続光下で 1 日置いた。1 日の光照射により、子葉は完全に緑化し、葉緑体もかなりのチラコイド膜を発達させる（図 11a)。Cpn10（AT3）の局在を示す金粒子は確かに葉緑体内に観察され、他のオルガネラには見られなかった。また、金粒子は全てチラコイド膜に接していることが観察された。コントロールとしてストロマタンパク質である RuBisCO についても同様に観察したところ、RuBisCO の局在を示す金粒子は葉緑体内に一様に分布していた（図 11b)。これらの分布の違いにより Cpn10（AT3）はチラコイド内腔に存在する可能性が示唆された。

<u>シロイヌナズナにおけるコ-シャペロニンの遺伝子発現</u>

シロイヌナズナより局在の異なるコ-シャペロニンホモログの cDNA が単離できたので、それぞれの遺伝子発現を比較した。まず、各コ-シャペロニンの植物体内における器官別の遺伝子発現をノザン解析により行った。ミトコンドリア局在型の Cpn10（AT1）は根、茎、葉のいずれの器官においても発現が見られた（図 12a)。Cpn20（AT2）も根では非常に発現が低いものの調べた全ての器

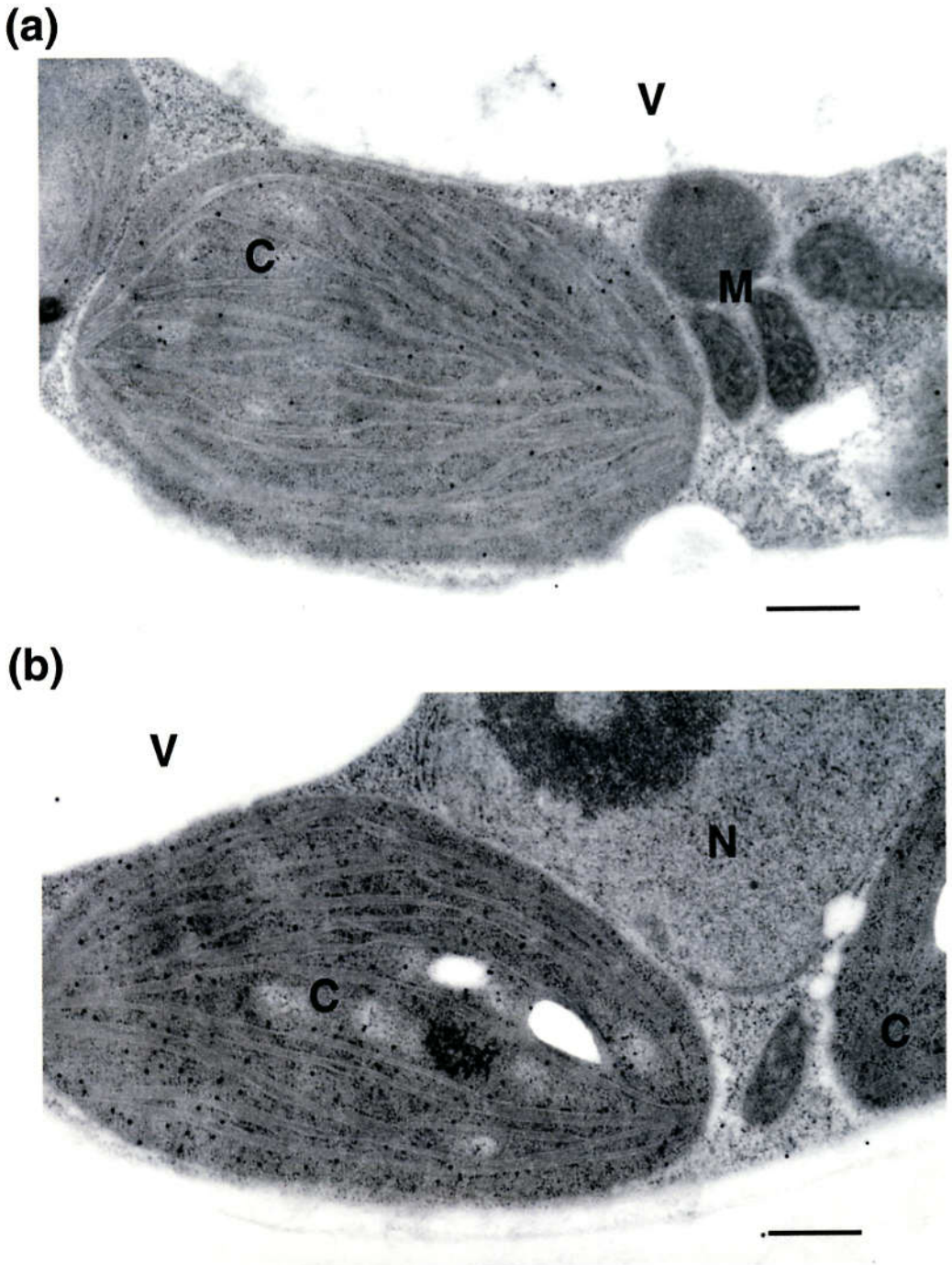


図11　Cpn10（AT3）を過剰発現する形質転換シロイヌナズナの免疫電顕
　形質転換シロイヌナズナを暗所、寒天培地上で発芽させ、暗所で4日間、その後連続光下で1日育てた。免疫電顕には子葉を用いた。His- Cpn10（AT3）に対する抗体を用いて行ったものを（a）に、ホウレンソウのRuBisCOに対する抗体を用いて行ったものを（b）に示した。バーの長さは0.5 μmで、Mはミトコンドリア、Cは葉緑体、Nは核、Vは液胞を示している。

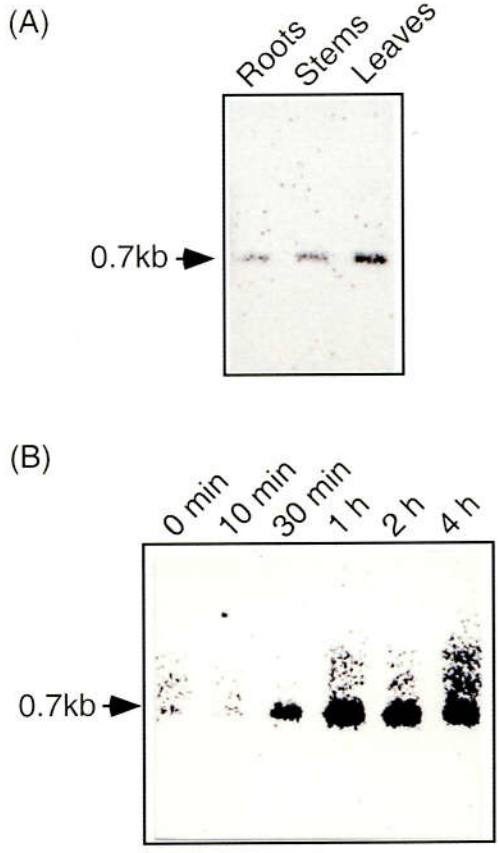


図12　シロイヌナズナにおけるCpn10（AT1）mRNA蓄積量の器官特異性（a）と熱処理に対する応答性（b）

（a）　根、茎、葉より抽出した全RNA10 μg分を電気泳動し、放射線ラベル　したCpn10（AT1）cDNAをプローブとしてノザン解析を行った。rRNAの量はほぼ等量であった。

（b）　シロイヌナズナの植物体に35℃の熱処理を行った。一定時間ごとに葉を切り取り全RNAを抽出し、（a）と同様にノザン解析を行った。熱処理時間は各レーンの上に示した。

官で発現していた（図 13a)。一方、Cpn10（AT3）は茎と葉のみで発現が見られ、根での発現は見られなかった（図 14a)。このことから Cpn10（AT1）と Cpn20（AT2）は各器官のミトコンドリアとプラスチドにおいて細胞の維持や活動に基本的な役割を果たしているのと考えられた。また Cpn10（AT3）は葉や茎といった緑色器官においてのみ発現がみられることから、特に葉緑体において機能する可能性が考えられる。

次に、各コ-シャペロニンの熱ストレスへの応答性を調べたるために、22℃で約 1 ヶ月間育てたシロイヌナズナの植物体を 35℃のインキュベーターに移して熱処理を行った。一定時間ごとに熱処理を加えた葉を切り取り、その葉より全 RNA を抽出して各コ-シャペロニンの発現量を調べた。ミトコンドリア局在型の Cpn10（AT1）は熱処理 30 分で既に mRNA の蓄積が見られ始め、1 時間までは蓄積量が増大し、その後調べた範囲の時間では蓄積量は変化なく高い発現が続くことがわかった。（図 12b)。Cpn20（AT2）も熱処理 30 分で mRNA の蓄積が見られ始めていたが、調べた時間内では蓄積量は増大し続けた（図 13b)。これらに対し、Cpn10（AT3）の mRNA の蓄積量は熱処理によりほとんど変動しなかった（図 14b)。さらに、このノザン解析の結果を確かめるため定量的 PCR も行った。まず、5 μg の全 RNA をオリゴ(dT)プライマーを用いて逆転写させた。コントロールとして APETARA2(AP2)の蓄積量も測定した。図 14c には AP2 で標準化した Cpn10（AT3）の蓄積量を示した。定量的 PCR の結果もノザンの結果と同様、熱処理により蓄積量の変動がないことを示していた。

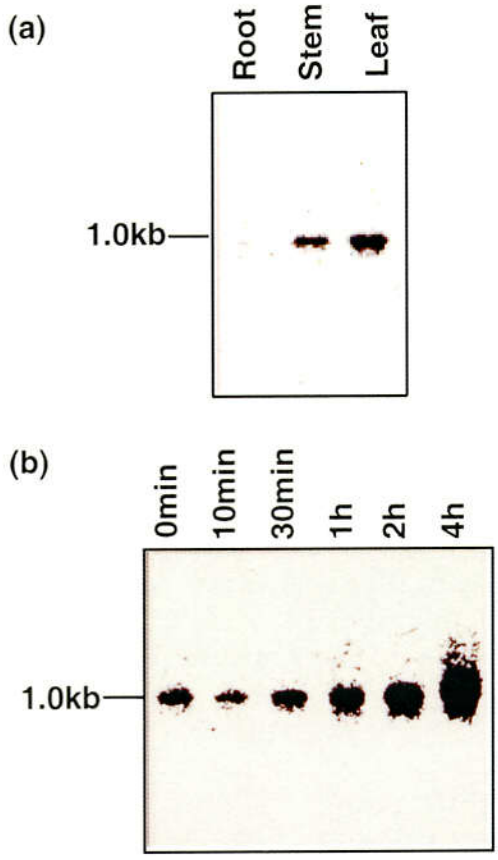


図13　シロイヌナズナにおけるCpn20（AT2）mRNA蓄積量の器官特異性（a）と熱処理に対する応答性（b）

(a)　根、茎、葉より抽出した全RNA10 μg分を電気泳動し、放射線ラベルしたCpn20（AT2）cDNAをプローブとしてノザン解析を行った。rRNAの量はほぼ等量であった。

(b)　シロイヌナズナの植物体に35℃の熱処理を行った。一定時間ごとに葉を切り取り全RNAを抽出し、（a）と同様にノザン解析を行った。熱処理時間は各レーンの上に示した。

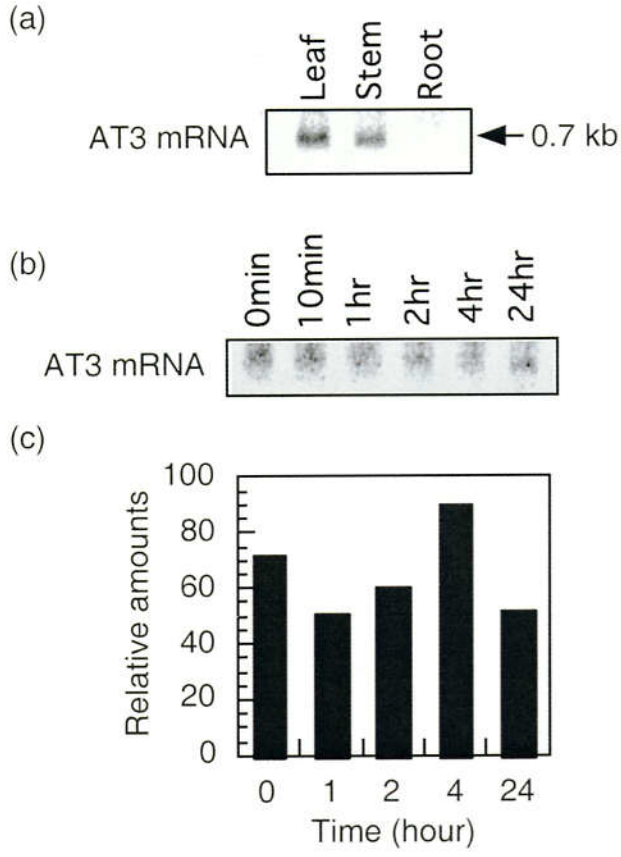


図14　シロイヌナズナにおけるCpn10（AT3）mRNA蓄積量の器官特異性（a）と熱処理に対する応答性（b）

（a）　根、茎、葉より抽出した全RNA20 μg分を電気泳動し、放射線ラベルしたCpn10（AT3）cDNAをプローブとしてノザン解析を行った。rRNAの量はほぼ等量であった。

（b, c）　シロイヌナズナの植物体に35℃の熱処理を行った。一定時間ごとに葉を切り取り全RNAを抽出し、ノザン解析（b）又は定量的PCR（c）を行った。AP2のmRNA量で標準化したCpn10（AT3）mRNAの量を棒グラフで示した。熱処理時間は各レーンの上又は下に示した。

In vitro におけるコ-シャペロニン活性の解析

各コ-シャペロニンの機能的な特徴を調べるために、シャペロニンを介したタンパク質巻き戻り実験を行った。各コ-シャペロニンはアミノ末端にヒスチジンタグ（His-tag）を付加して大腸菌で発現させたものを実験に用いた。延長配列を持つ Cpn20（AT2）、Cpn10（AT3）は成熟部分のみを発現させた。His-Cpn10（AT1）と His-Cpn20（AT2）は大腸菌内で大量に可溶性タンパク質として発現したので、それぞれ超音波破砕した大腸菌の上清を Ni を結合させたアフィニティーカラム（Ni カラム）にかけて精製した。His-Cpn10（AT3）は不溶性の封入体となったので、まず 8 M 尿素で可溶化してから Ni カラムに吸着させた。その後、溶出液の尿素濃度を 6 M から 1 M に下げるグラジエントをかけ、カラム上で His-Cpn10（AT3）の再生（巻き戻り）を行った。以上の操作により可溶化した His-Cpn10（AT3）をイミダゾールで溶出して実験に用いた。各溶出画分を SDS-PAGE にかけ CBB 染色を行った結果、どのコ-シャペロニンも単一バンドを示し、十分精製されていることがわかった（図 15a）。

巻き戻り実験には基質としてクエン酸合成酵素（Citrate synthase, CS）を用い、相手方のシャペロニンとしては大腸菌の GroEL を用いた。ポジティブコントロールとして GroEL の本来のコ-シャペロニンである GroES で同様の実験を行ったところ、変性させた CS の約 75%の活性を回復させることができた（図 15b）。His-Cpn10（AT1）は GroES と同程度の活性を示した。これによりアミノ末端のヒスチジンタグよる影響はほとんどないことがわかった。His-Cpn20（AT2）は同様の条件下で CS の約 45%の活性を回復させることができた。His-Cpn20（AT2）の量を倍にしても変化はなかったので、コ-シャペロニンの量としては

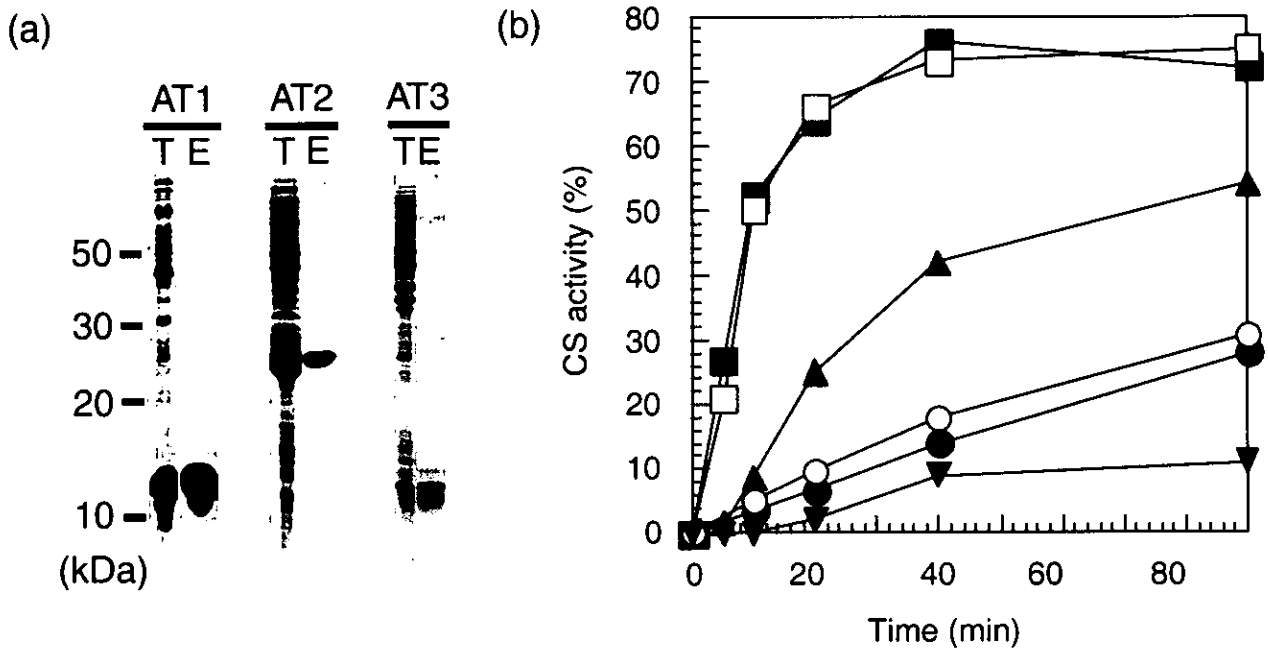


図15　GroELを用いたin vitro巻き戻り実験におけるシロイヌナズナのコーシャペロニンの活性

(a)　各コーシャペロニンはアミノ末端にヒスチジンタグを付加した形で大腸菌で発現させたものを用いた。各コーシャペロニンを発現している大腸菌の全タンパク質（T）と精製タンパク質（E）をSDS-PAGEにかけCBB染色した像を示した。

(b)　巻き戻り実験の基質としてクエン酸合成酵素（CS）、シャペロニンとしてGroELを用いた。変性させたCSは0℃でGroELの14量体（GroEL14）を含む溶液で150 nMになるように希釈した。その後、2 mM ATPとコーシャペロニンを加え35℃で反応させることにより、GroELを介してCSを巻き戻らせた。CSの活性の回復をコーシャペロニンの活性の指標として測定した。加えたコーシャペロニンの種類と濃度は300 nM GroES7量体（■）、300 nM His-Cpn10（AT1）7量体（□）、300 nM His-Cpn20（AT2）4量体（▲）、600 nM His-Cpn10（AT3）7量体（○）、300 nM His-Cpn10（AT3）7量体（●）、バッファーのみ（▼）である。

十分量含まれていると考えられる。従って、相手方として用いたGroELに対して His-Cpn20（AT2）の比活性が低いことが示唆された。His-Cpn10（AT3）ではさらに比活性が低くなることがわかった。His-Cpn10（AT3）はタンパク質を調製する際に再生操作をおこなっているため正しい高次構造をとっていない可能性がある。または、配列上非常に特異性が高いためGroELとの相互作用が弱くなっているかもしれない。

His-Cpn20（AT2）の多量体構造

ゲルろ過カラムクロマトグラフィーによる解析

GroESは7量体で機能することが知られているが、Cpn20（AT2）はGroES様ドメインを2つ持つため同様の多量体構造をとることはできない。Cpn20（AT2）の多量体構造を調べるため、His-Cpn20（AT2）を用いてゲルろ過カラムクロマトグラフィーを行った。His-Cpn20（AT2）の多量体もおそらくGroESの多量体と同様ドーム状の構造をとると予想され、ゲルろ過カラムでの挙動は球状タンパク質とは異なることが考えられる。そのため以下に示す計算分子量はあくまで比較のための数値として扱う。

Niカラムにより精製したHis-Cpn20（AT2）と、コントロールとしてHis-Cpn10（AT1）をゲルろ過カラムにかけ、280 nmの吸光度を測定した（図16a）。両タンパク質の溶出プロファイルの34分あたりに共通に見られるピークは、夾雑物として含まれる大腸菌由来の低分子量のタンパク質を示すと考えられる。His-Cpn20（AT2）をカラムにかけると、主なピークが25分あたりに、もう1

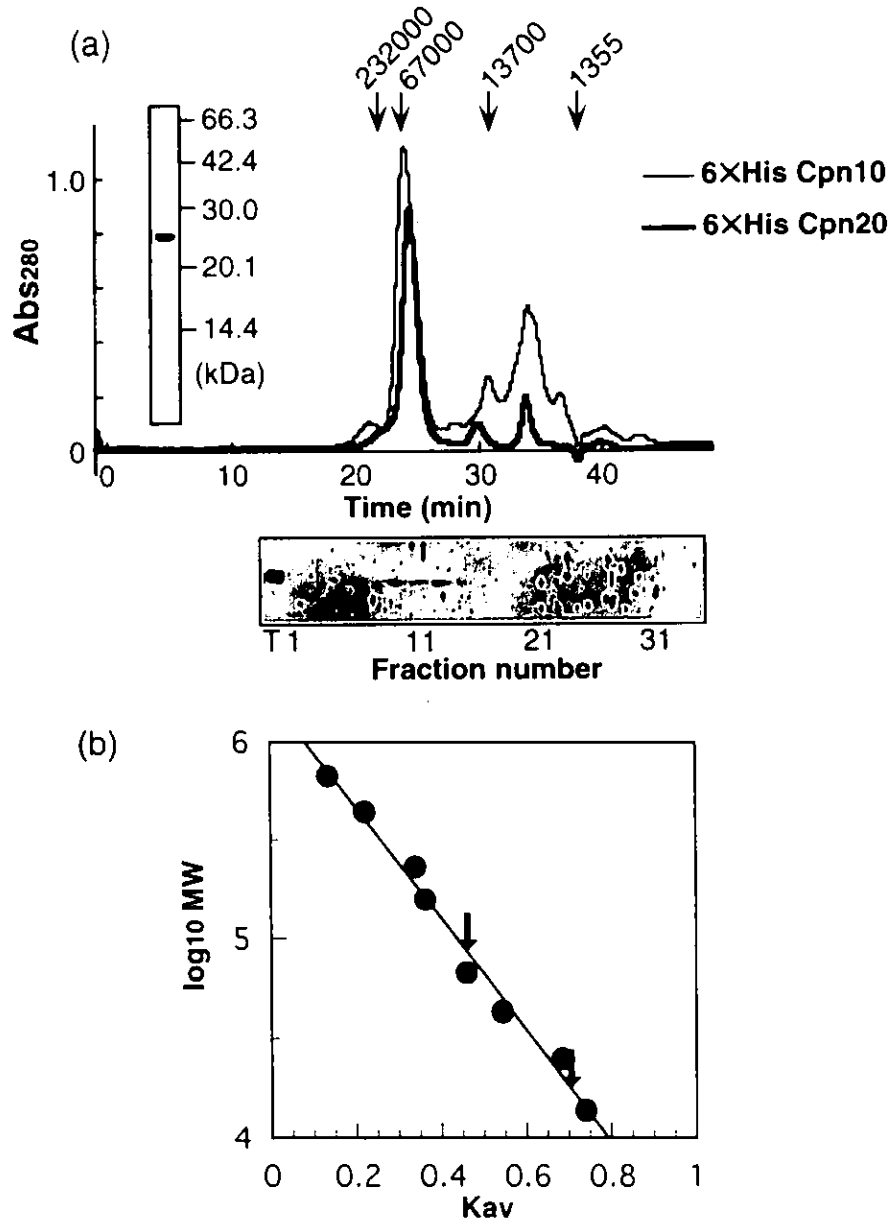


図 16　ゲルろ過カラムクロマトグラフィーによる His-Cpn20（AT2）の多量体の分子量の測定

(a)　Superose 12 カラムを用いた His-Cpn20（AT2）と His-Cpn10（AT1）のゲルろ過カラムクロマトグラフィーを行い、280 nm の吸光度を測定した（上）。使用した His-Cpn20(AT2)を SDS-PAGE し、CBB 染色した像を図中に挿入してある。太線は His-Cpn20 (AT2) を、細線は His-Cpn10 (AT1) を示す。分子量マーカーの溶出位置を分子量とともに矢印で示した。タバコの葉緑体の可溶性タンパク質を同じカラムにかけ、その溶出液を分画、回収した。各画分を SDS-PAGE にかけ His-Cpn20（AT2）に対する抗体を用いてイムノブロット解析を行った（下）。葉緑体の全タンパク質はレーンの一番左に示した。

(b)　HiLoad Superdex 200pg カラムを用いた場合の分子量マーカーのキャリブレーション曲線を示した。キャリブレーション曲線（y=-2.81x+6.23）は最小二乗法を用いて計算した。His-Cpn20（AT2）の Kav 値（0.461 と 0.705）は矢印で示した。

つのピークが 30 分あたりに見られた。これらは分子量 75000 と 16000 に相当していた。さらに詳しく調べるため、2 倍の長さのゲルろ過カラムを用いて行った結果、分子量 86000 と 18000 に相当するピークが見られた（図 16b)。これらの結果から、どちらのカラムにおいても His-Cpn20（AT2）は 4 量体または 5 量体であると推測された。コントロールとして行った His-Cpn10（AT1）については分子量 84000 と 12000 に相当するピークが見られ、予想通り 7 量体であるという結果を得ることができた。

内在性の Cpn20（AT2）について調べるため、タバコより葉緑体を精製しゲルろ過カラムにかけた。溶出液を分画して回収しイムノブロット解析を行った結果、His-Cpn20（AT2）の多量体の溶出位置と同じところにバンドが検出された（図 16a)。このことより、大腸菌で発現させた His-Cpn20（AT2）の構造は、内在性の Cpn20（AT2）と同様のものであることが示唆された。

架橋させた His-Cpn20（AT2）多量体の解析

His-Cpn20（AT2）が 4 量体か 5 量体かを明らかにするため、DSP を用いて架橋した多量体を調製し SDS-PAGE で調べた（図 17a)。DSP はタンパク質の末端のα-アミノ基とリジンのε-アミノ基に共有結合する。DSP を 500 から 1000 倍の濃度比で加えると、架橋された大きな多量体が観察された。架橋された多量体のうち最大のものは His-Cpn20（AT2）では 4 量体だった。架橋された His-Cpn20（AT2）4 量体の SDS-PAGE 上での移動度は His-Cpn10（AT1）7 量体の移動度よりやや大きく、ゲルろ過カラムクロマトグラフィーによる解析の結果と一致していた。高濃度で DSP を加えることによりタンパク質が凝集してしまうので、

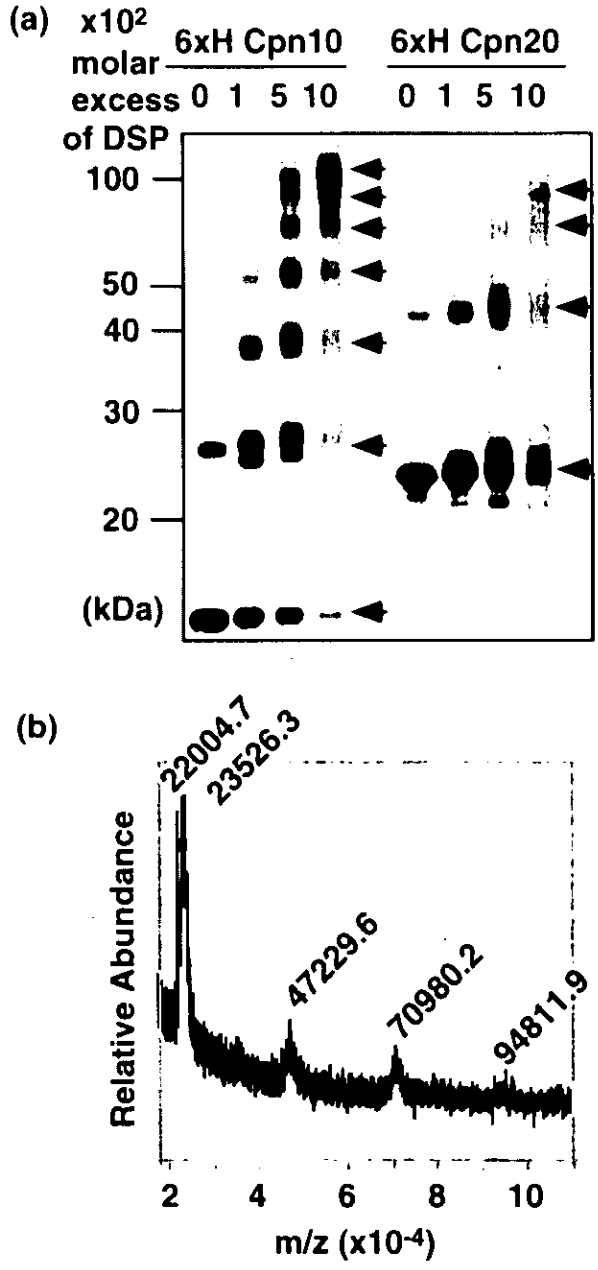


図 17　架橋剤 DSP を用いた His-Cpn20（AT2）の多量体構造の解析

(a)　種々の濃度の DSP と反応させた His-Cpn20（AT2）を右に、His-Cpn10（AT1）を左に示した。SDS-PAGE は還元剤（2ME）を除いて行い、CBB 染色した。単量体と多量体の位置を矢印で示した。分子量マーカーの位置は左に示した。

(b)　1000 倍濃度の DSP で架橋させた His-Cpn20（AT2）を質量分析した。測定された分子量はそれぞれのピークの上に示した。

架橋された 4 量体のみを調製することはできなかった。

次に、架橋した種々の多量体が混在する状態で質量分析を行った。図 17b に示した通り 94811.9、70980.2、47229.6、23526.3 Da のピークが観察され、SDS-PAGE 上で見られた 4 本のバンドがそれぞれ単量体から 4 量体に相当するものであることが確認された。4 量体に相当するシグナルは非常に小さいが、これは分子量の大きいものほどイオン化の効率が悪いためである。His-Cpn20（AT2）の塩基配列より計算された分子量 22031 Da と単量体の観察された分子量 23526.3 Da の比較より、単量体の His-Cpn20（AT2）1 分子あたり DSP が 4 分子が結合していると考えられる。

非変性 PAGE による His-Cpn20（AT2）多量体の解析

His-Cpn20（AT2）が 4 量体であることを確認するために、非変性 PAGE により His-Cpn20（AT2）多量体の分子量を決定した。未変性のタンパク質を SDS を含まない濃度の異なるアクリルアミドゲル（6～10%）を用いて 4℃で泳動し、Hedrick と Smith の方法（1968）に従いゲル濃度の変化に伴うタンパク質の移動度の変化により分子量を決定した。まず、BPB の移動度に対する各タンパク質の相対的な移動度（Relative mobility, Rm）を求め、その対数をゲル濃度に対してプロットすることにより、各タンパク質に固有の傾きを持つ直線を引くことができる。その傾きはタンパク質の分子量と相関関係を示す（図 18）。His-Cpn20（AT2）多量体の傾きは 8.6×10^{-2} であり、その傾きから計算された分子量は 90 kDa であった。この値（90）を His-Cpn20（AT2）の計算分子量 22 kDa で割ると 4.1 という値が得られ、この手法からも His-Cpn20（AT2）多量体が 4 量体であるこ

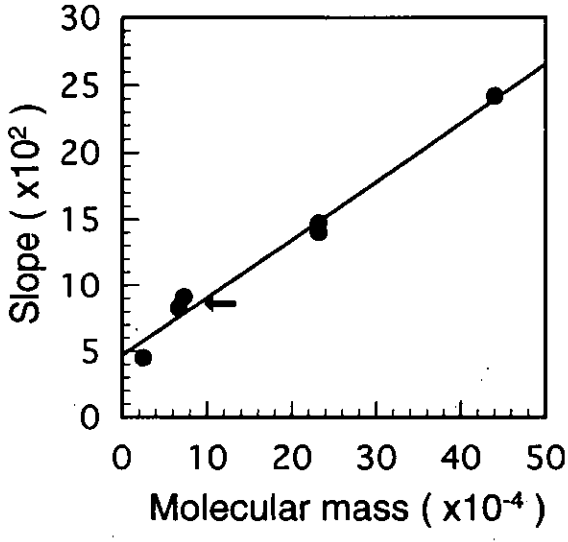


図 18　非変性 PAGE による His-Cpn20（AT2）多量体の分子量の測定

分子量マーカーとして、キモトリプシノーゲン（分子量 25 kDa）、BSA（67 kDa）、GroES（73 kDa）、カタラーゼ（232 kDa）、フェリチン（440 kDa）を各 5 μg ずつ用いた。非変性 PAGE は 6～10%まで 1%ずつアクリルアミド濃度を変えて行った。傾き（Slope）は、ゲル濃度に対する各分子量マーカーの相対的な移動度をプロットし計算した。分子量に対して傾きをプロットし、キャリブレーション曲線を作成した。キャリブレーション曲線（y=0.438x+4.66）は最小二乗法を用いて計算した。His-Cpn20（AT2）の傾き（8.6）は矢印で示した。

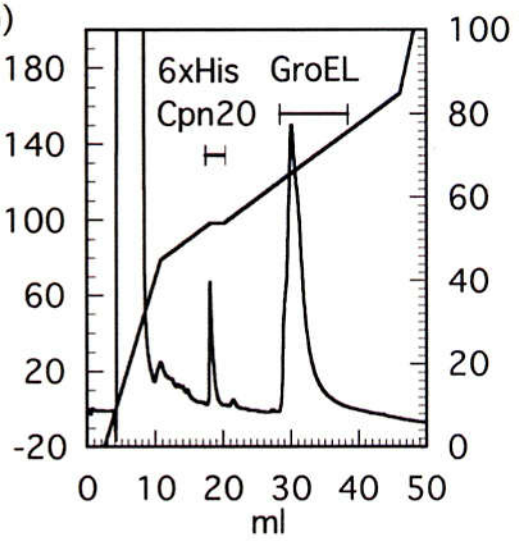
とが確認された。

His-Cpn20（AT2）とGroELの相互作用

His-Cpn20（AT2）が4量体であるとGroES様ドメインを8個含むことになり、本来のGroESの7量体構造とは異なる。しかしながら、His-Cpn20(AT2)はGroESと同様に、GroELと相互作用しコ-シャペロニンとして機能できることがわかっている（図15)。そこで、複合体に含まれるHis-Cpn20（AT2）とGroELのそれぞれの量比を調べた。His-Cpn20（AT2）とGroELの複合体は1 mM ADP中で安定に存在するので、ADP 共存下で複合体を調製した。調製の各過程を図 18に示した。His-Cpn20（AT2）とGroELの混合液（図19aのレーン1）をゲルろ過カラムにかけることにより、複合体の含まれる画分（図19aのレーン3と4）と結合していないHis-Cpn20（図 19aのレーン7〜9）に分離する。今回の実験条件では結合していないHis-Cpn20（AT2）はほとんど溶出していなかった。次に複合体の含まれる画分を Ni カラムにかけた。His-Cpn20（AT2）と結合しているGroELのみが吸着しイミダゾールで溶出された（図19aのレーンa〜c)。吸着しなかったGroELは複合体を形成していない（図 19aのレーンF)。複合体は逆相カラムでそれぞれのタンパク質ごとに回収し（図 19b)、アミノ酸分析を行った。回収したタンパク質のうち、His-Cpn20（AT2）は20%、GroELは2%分を分析に用いた。定量性の高い3種類のアミノ酸について図19cに結果を示した。観察されたアミノ酸量（Amount）を各アミノ酸の含有量（Residue number）で割り、タンパク質量（Protein）を計算した。複合体に含まれていたHis-Cpn20（AT2）とGroELの量はそれぞれ0.32 nmolと1.26 nmolとなり、その比は3.6：

(a)

(b)



(c)

GroEL

Amino acid	Amount (nmol)	Residue number	Protein (nmol)
Ala	92.95	74	1.26
Leu	52.55	42	1.25
Phe	8.90	7	1.27

6xHis Cpn20

Amino acid	Amount (nmol)	Residue number	Protein (nmol)
Ala	4.33	13	0.33
Leu	5.72	18	0.32
Phe	0.94	3	0.31

図 19　GroEL と His-Cpn20（AT2）複合体に含まれる各タンパク質の量比の決定

(a)　GroEL と His-Cpn20(AT2)複合体は 1 mM ADP 存在下で調製した。GroEL の位置を黒丸で、His-Cpn20（AT2）の位置を矢印で示した（レーン a)。精製前の混合液（レーン 1）とゲルろ過カラムクロマトグラフィー後の画分（レーン 2〜9）及び、複合体を含む画分（レーン 3 と 4）を Ni カラムにかけたた時の非吸着画分（レーン F)、洗浄画分（レーン W）と溶出画分（レーン a〜e）を SDS-PAGE にかけ CBB 染色した。10 kDa のラダーマーカーをレーン M に示した。マーカーの一番下のバンドは 20 kDa に相当する。

(b)　GroEL と His-Cpn20（AT2）は精製した複合体を逆相カラムにかけて分離した。220 nm の吸光度（mAU220)、溶出に用いたアセトニトリル の濃度（AN）を示した。 GroEL と His-Cpn20（AT2）それぞれの溶出位置は線で示した。

(c)　アラニン、ロイシン、フェニルアラニンについてアミノ酸分析の結果を示した。分析した量は GroEL が 2%、His-Cpn20（AT2）は 20%だったので、それから計算した全量（Amount）を 2 列目に示した。残基数（Residue number）はアミノ酸配列から計算した。これらの値より計算した各タンパク質の量を一番右に示した。

14 となった。この結果から GroEL14 量体と His-Cpn20（AT2）4 量体が複合体を形成していることが明らかとなった。

葉緑体に局在するコ-シャペロニン

2 種類のコ-シャペロニン

葉緑体には Cpn20（AT2）と Cpn10（AT3）の 2 種類のコ-シャペロニンが存在しており（図 8 と 10）、特に Cpn10（AT3）は葉緑体に特異的であることが発現パターンから示唆された（図 14）。これら 2 種類のコ-シャペロニンについてさらに調べるために、子葉の発芽と緑化に伴うタンパク質蓄積量の変動を調べた。まず、シロイヌナズナの種子を暗所で発芽させた。子葉は暗所で生育している間は白色であるが、連続光を 1 日照射した後では完全に緑化している。光照射により、12S グロブリンなどの貯蔵タンパク質は急速に分解され、LHCII などの光合成の電子伝達に関わるタンパク質が蓄積される（図 20）。RuBisCO のような炭酸固定に関わるタンパク質は暗所でも 5 日をすぎると蓄積が観察されるようになる。また、暗所 5 日（図 20 レーン 5）と暗所 4 日明所 1 日（図 20 レーン 4D1L）の比較から、光による発現誘導も確認された。イムノブロット解析により Cpn20（AT2）と Cpn10（AT3）の蓄積量を調べた結果、Cpn20（AT2）は発芽の初期から発現が見られたのに対し、Cpn10（AT3）は暗所 5 日ではわずかに蓄積が見られる程度であった（図 20）。Cpn10（AT3）の蓄積は他の葉緑体のタンパク質と同様、光照射により増加した。以上の結果は、Cpn20（AT2）がプラスチドの基本的な機能に、Cpn10（AT3）が葉緑体の機能に何らかの働きを

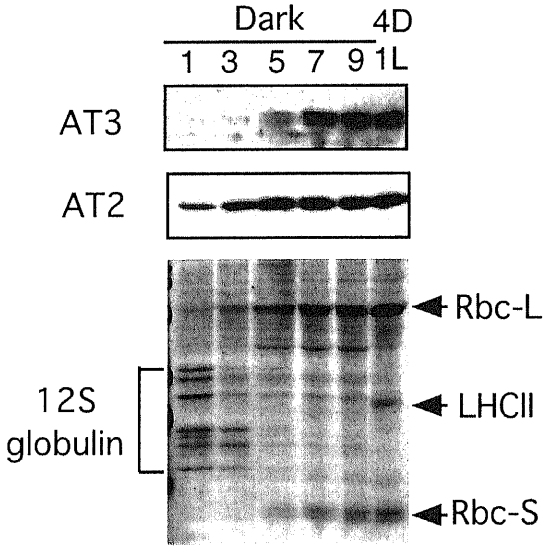


図20シロイヌナズナ子葉における発芽および緑化に伴うタンパク質蓄積量の変化

種々の条件で生育させた子葉より抽出したタンパク質（20 μg）をSDS-PAGEにかけイムノブロット解析を行った。His-Cpn10（AT3）に対する抗体を用いた場合を上に、His-Cpn20（AT2）に対する抗体を用いた場合を真ん中に示した。レーンの上に示した数字は暗所での生育日数を示す。4D1Lは4日間暗所で育てた後、1日連続光を照射した子葉を示す。イムノブロットに使用したPVDF膜をCBB染色したものを下に示した。

もつことを支持するものとなった。

考察

シロイヌナズナのコ-シャペロニンホモログの cDNA クローニング

シャペロニンは多量体で大きな円筒形の構造を形成し、その内部をタンパク質の巻き戻りの‘場’として提供している（Braig ら、1994）シャペロニンは細胞の構築、維持に必須なタンパク質であり、原核細胞から真核細胞まで種々の生物が共通に持つタンパク質である。1 次配列上からはシャペロニンは 2 つのグループに分類され、それぞれ進化上関連深い場所に存在する（Braig、1998）。高等植物の細胞内で働くシャペロニンの全体像を解明するため、まずグループ 1 のシャペロニンの制御因子であるコ-シャペロニンについて、シロイヌナズナを用いて解析した。

コ-シャペロニンの cDNA クローニングのため、大腸菌のコ-シャペロニンホモログである GroES の温度感受性変異体を相補する cDNA をスクリーニングした（図 3）。得られたクローンは 10 kDa のコ-シャペロニンホモログ（Cpn10、AT1 と呼ぶ）で、ミトコンドリアに局在するホモログであった。AT1 は酵母（Rospert ら、1993）やほ乳類（Ryan ら、1994）のミトコンドリア Cpn10 と同様に延長配列を持っていない（図 4）。そのため大腸菌で成熟型として発現し、GroES の機能を相補できたと考えられる。実際に in vitro での活性を調べた結果、AT1 は GroES と同様の活性を持つことが確かめられた（図 15）。EST クローンを利用することにより、葉緑体のホモログの cDNA を 2 種類単離することができた（図 7 と 9）。1 つは Cpn10（AT3 と呼ぶ）であったが既知のホモログとは非常に相

同性が低く、もう 1 つは Cpn10（GroES）様ドメインを 2 つ持つ Cpn20（AT2 と呼ぶ）であった。どちらも約 50 残基の延長配列を持つ前駆体として合成されることがわかった。成熟型のアミノ末端付近にメチオニンを持っていれば、上述のスクリーニングにおいて部分的に発現した cDNA クローンを単離できる可能性も考えられる。しかし、AT2、 AT3 ともその条件を満たしておらず、延長配列を持つことが上述のスクリーニングで単離できなかった原因の 1 つと考えられる。また、大腸菌のシャペロニンホモログである GroEL を調節する in vitro での活性を測定したところ、GroES に比べ AT2 で 80%、AT3 で 40%と低いことがわかった（図 15）。この活性の差は GroEL との相互作用の強さに関係していると考えられる。シャペロニンの研究は主として大腸菌の GroEL、GroES について行われているが、葉緑体のホモログはそれらとは異なる興味深い特徴があると考えられる。

<u>葉緑体の 2 量体型コ-シャペロニン</u>

葉緑体の 2 種類のコ-シャペロニンの内、特に注目される特徴は 2 量体型コ-シャペロニン（Cpn20、AT2）の多量体構造である。なぜなら、GroES が 7 量体であることは結晶構造解析からも明らかだが（Hunt ら、1996）、AT2 はそれ自身 2 量体であるので GroES の 7 量体と同等の構造をとり得ないからである。AT2 の本来の相手方のシャペロニンが 7 量体でなく、6 又は 8 量体を形成している可能性も考えられる。ただ、AT2 は GroEL とも相互作用できることが示されており（図 15、Baneyx ら、1995、Bertsch と Soll、1995）、シャペロニンとコ-シ

ャペロニンの多量体構造が必ずしも一致している必要はないと推測される。そこで、AT2 を大腸菌でヒスチジンタグをアミノ末端に付加した形で発現させ、その精製産物（His-AT2）を用いて Cpn20 の多量体構造を解析した。

まず、ゲルろ過カラムクロマトグラフィーを行って単量体、多量体それぞれの分子量を測定した（図 16)。測定した値を単純にわり算してみると、30 cm のカラム（Superose 12、分画範囲 1×10^3～3×10^5）においては 75/16=4.7、60 cm のカラム（Superdex 200pg、分画範囲 1×10^4～6×10^5）においては 86/19=4.8 となった。非変性 SDS-PAGE による測定では、多量体の大きさは 90 kDa となり、単量体の計算分子量 22 kDa でわると 4.1 となった（図 18)。架橋した多量体の中には 5 量体は観察されなかったので（図 17)、His-AT2 は 4 量体であると結論した。また、GroEL と His-AT2 の複合体を調製し、複合体に含まれる量比を調べたところ、14 : 3.6 という比であった（図 19)。これは GroEL の 14 量体と His-AT2 の 4 量体が結合していることを強く示唆する結果となった。

では、シャペロニンとコ-シャペロニンはどのように結合しているのだろうか。GroEL と GroES の場合、GroES の持つモバイルループと呼ばれるβヘアピン構造が GroEL との相互作用に関わっている（Landry ら、1993、Xu ら、1997)。このモバイルループと GroEL の相互作用は基質であるほどけたポリペプチドと GroEL の相互作用を分断し、ポリペプチドの巻き戻りを促進させると考えられている。基本的に GroES のモバイルループと GroEL の結合は 1 対 1 であり、GroES と GroEL の複合体には 7 カ所で同様の結合が見られるはずである。最近、GroEL の 7 量体を 1 本のポリペプチド鎖として発現させることで、任意の変異体を任意の位置に導入する試みがなされた（Farr ら、2000)。この実験により、7 カ所

のうち半分程度の結合部位が正常であれば、十分機能を果たすことが示された。この時、野生型 GroEL は散在するより 1 カ所にまとまって存在する方が機能性が高いことも示された。このことから、AT2 も 8 本のモバイルループ全部が GroEL と結合しなくても十分に制御因子として活性を持つことができると考えられる。

葉緑体での AT2 の本来の相手方は RuBisCO 結合タンパク質として同定されたシャペロニン（Cpn60αと Cpn60β）である。この 2 種のタンパク質は相同性が 50%程度であり、ヘテロな多量体を形成している（Nishio ら、1999)。Cpn60α は単独では多量体構造を形成できず、Cpn60βの共存下でのみヘテロな多量体を形成することができる（Clony ら、1992、Dickson ら、2000)。一方、Cpn60βは単独でも多量体を形成することができ、Cpn60βのホモ多量体はミトコンドリア Cpn10と協同して機能し得るが、葉緑体 Cpn20 や GroES とは協同して機能できないことが示された（Dickson ら、2000)。このように葉緑体のシャペロニンシステムはコ-シャペロニンだけでなくシャペロニン自体もヘテロな性質を持っており、GroEL と GroES の研究からだけでは推測できない部分がある。ただ、Cpn60αと Cpn60βの多量体構造は GroEL の多量体構造とよく似ていることから、サブユニット数は GroEL と同じ 14 であると考えられる。シャペロニンと AT2 との結合様式については今後さらに解析が必要である。

<u>葉緑体に局在するもう 1 つのコ-シャペロニン</u>

ミトコンドリアのシャペロニンシステムが大腸菌の GroEL と GroES と非常に

類似していたのに対し，葉緑体のシステムはタンパク質の特徴としては似ているものの、いくつかの点で大腸菌のものとは異なっていることがわかった。葉緑体の起源とされるシアノバクテリアには Cpn20 のホモログはなく、GroEL、GroES それぞれのホモログのみが見られることから、葉緑体のシャペロニンシステムは共生後に大きく変化したと思われる。Schlicher と Soll（1996）はコ-シャペロニンに対するペプチド抗体によるイムノブロット解析より、葉緑体のチラコイド内腔に GroES（Cpn10）型のコ-シャペロニンが存在することを示唆した。EST データベースの活用により、葉緑体には確かに GroES（Cpn10）型のコ-シャペロニンが存在することがわかった（図 9 と 10)。このホモログ（AT3 と呼ぶ）の局在性についてはさらに検討が必要であるが、免疫電顕像からはチラコイド膜付近に金粒子が見られることからチラコイド内腔に存在することが示唆される。AT3 の発現パターンは葉緑体の発達、特にチラコイド膜構造の発達と関連していることから、チラコイド膜に多く含まれる光合成に関わるタンパク質が基質である可能性も考えられる。シャペロニンについても EST データベースを検索してみると、既知の Cpn60αと Cpn60β以外にもトランジットペプチドを持つシャペロニンがいくつか存在することが示唆された。これらは Cpn60αより Cpn60βに似ていることから AT3 の相手方が Cpn60β様タンパク質である可能性がある。Cpn60βは単独で多量体形成が可能なことから、Cpn60β様タンパク質も単独で機能できると考えられる。従って、葉緑体にはヘテロな多量体を形成するシャペロニンシステムに加え、AT3 と Cpn60（γ?）というホモ型多量体同士で構成されているシャペロニンシステムが存在するのではないかと推測される。

既知のコ-シャペロニンホモログの系統樹を作成したところ、AT3 はどのホモログとも相同性が低いことがわかった（図 11)。詳細に調べるため、結晶構造解析のされている GroES と比較した（図 9b)。GroES は 9 つのβシートとモバイルループより構成されている（Hunt ら、1996)。4 番目と 5 番目のβシートによってルーフと呼ばれる構造が形成されている。GroES の 7 量体はドーム上の構造をとるのだが、名前通り、屋根（天井）に当たる部分をルーフ構造を呼ぶ。AT3 のリジン 106 とアルギニン 107 の間に 8 残基のギャップを挿入し、GroES のモバイルループに 2 残基のギャップを挿入することにより、3、6, 9 番目のβシートの同一性が増加した（図 9b)。このような比較を行ってみると、AT3 はルーフ構造に当たる領域が大きく変化したコ-シャペロニンであると推測される。ルーフ構造が完全に欠如したタイプのコ-シャペロニンとしては T4 ファージのタンパク質である Gp31 が知られている。Gp31 は分子量も GroES よりやや大きくアミノ酸配列の相同性も低いが、結晶構造解析よりコ-シャペロニンと非常によく似た構造をとっていることが示された（Hunt ら、1997)。T4 ファージの主要なキャプシドタンパク質である Gp23 の高次構造形成には、シャペロニンとして宿主の GroEL を利用するのだが、Gp31 の欠損した T4 ファージではそれが起こらず GroES では代用がきかない（Laemmli ら、1970)。このことから、Gp31 のルーフ構造が欠如しているということが GroEL 内で Gp23 の巻き戻りを進めるために重要であると考えられた。Gp23 が 55 kDa と比較的大きいタンパク質であることから、より大きな基質に対応するために進化した結果であると想像されている。AT3 も同様な理由によりルーフ構造に変異が見られるとすると、AT3 の特異的な基質が存在する可能性が高いと考えられる。AT3 の機能を抑え

た植物体を作製することにより、そのような基質を特定できるのではないかと考えている。

その他のシャペロニンシステム

シロイヌナズナのコ-シャペロニンを調べることにより、葉緑体に存在するシャペロニンシステムの全体像が見えてきた。動物ではミトコンドリアのタンパク質がミトコンドリア以外の場所で機能するという報告がいくつかあり（Soltys と Gupta、1999）、シャペロニン（Hsp60）の細胞質における機能の報告もある。また、初期妊娠因子（Early pregnancy factor）として働く分泌タンパク質の実体がミトコンドリア Cpn10 そのものであったという驚くべき発見も報告された。それぞれのシャペロニン、コ-シャペロニンが多様な機能を示す可能性が提起されたといえる。

最近、グループ 2 のシャペロニンと協調して働く低分子量のタンパク質（GimC 又はプレフォルディン、PFD）の存在が明らかになってきた（Vainberg ら、1998）。GimC/PFD はコ-シャペロニンと同様多量体を形成して機能する。しかし、それ自身にペプチドと結合する能力を持ちシャペロニンとの結合も一時的な弱いものであるため、コ-シャペロニンではなく Hsp70 に似た機能を示すことがわかった。現在のところ、古細菌（Leroux ら、1999）、酵母（Geiser ら、1997）、哺乳動物（Vainberg ら、1998）で見つかっている。ホモロジー検索を行うと、シロイヌナズナや小麦の EST データベースにホモログが存在しており、植物のグループ 2 のシャペロニンシステムは他の真核生物のものと顕著な違いはないだ

ろうと推測される。このように、植物細胞におけるシャペロニンシステムは、他の真核生物と同様、真生細菌由来の GroEL-GroES 型と古細菌由来の細胞質型の 2 つであることが確かめられた。ただ、葉緑体においては、本来の GroEL-GroES 型と異なる特徴を示し、おそらく基質となるタンパク質の特異性に合わせて変化してきたのだろうと考えられる。これから in vivo での働きを調べることによりシャペロニンシステムの具体的な基質となるタンパク質が同定されれば、葉緑体におけるシャペロニンシステムの役割が明らかになると期待される。

引用文献

Anfinsen, C.B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230.

Arnon, D.I. (1949) Copper enzymes in chloroplasts. Polyphenol oxidase in *Beta vulgaris*. *Plant Physiol.* **24**, 1-15.

Asahi, M., Fujii, J., Takao, T., Kuzuya, T., Hori, M., Shimonishi, Y. and Taniguchi, N. (1997) The oxidation of selenocysteine is involved in the inactivation of glutathione peroxidase by nitric oxide donor. *J. Biol. Chem.* **272**, 19152-19157.

Baneyx, F., Bertsch, U., Kalbach, C.E., Vies, S.M.v.d., Soll, J. and Gatenby, A.A. (1995) Spinach chloroplast cpn21 co-chaperonin possesses two functional domains fused together in a toroidal structure and exhibits nucleotide-dependent binding to plastid chaperonin 60. *J. Biol. Chem.* **270**, 10695-10702.

Bertsch, U. and Soll, J. (1995) Functional analysis of isolated cpn10 domains and conserved amino acid residues in spinach chloroplast co-chaperonin by site-directed mutagenesis. *Plant Mol. Biol.* **29**, 1039-1055.

Bertsch, U., Soll, J., Seetharam, R.and Viitanen, P.V. (1992) Identification, characterization, and cDNA sequence of a functional 'double' groES-like chaperonin from chloroplasts of higher plants. *Proc. Natl. Acad. Sci.* **89**,8696-8700.

Bethtold, N., Ellis, J. & Pelletier, G. (1993) In planta Agrobacterium mediated gene transfer by infiltration of adult arabidopsis thaliana plants. *C.R.Acad. Sci.Ser. III Sci. Vie.* **316**, 1194-1199.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L. and Sigler, P.B. (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* **371**, 578-586.

Braig, K. (1998) Chaperonins. *Curr. Opin. Stuctr. Biol.* **8**, 159-165.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81**, 1991-1995.

Cloney, L.P., Wu, H.B. and Hemmingsen, S.M. (1992) Expression of plant chaperonin-

60 genes in *Escherichia coli*. *J. Biol. Chem.* **267**, 23327-23332.

Dickson, R., Weiss, C., Howard, R.J., Alldrick, S.P.,Ellis, R.J., Lorimer, G., Azem, A. and Viitanen, P.V. (2000) Reconstitution of higher plant chloroplast chaperonin 60 tetradecamers active in protein folding. *J.Biol.Chem.* **275**, 11829-11835.

Dubaquié, Y., Looser, R., Fünfschilling, U., Jenö, P. and Rospert, S. (1998) Identification of *in vivo* substrates of the yeast mitochondrial chaperonins reveals overlapping but non-identical requirement for hsp60 and hsp10. *EMBO J.* **20**, 5868-5876.

Ellis, R.J. and var der Vies, S.M. (1991) Molecular chaperones. *Annu. Rev. Biochem.* **60**, 321-347.

Ewalt, K. L., Hendrick, J., P., Houry, W., A. and Hartl F., U. (1997) In vivo observation polypeptide flux through the bacterial chaperonin system. *Cell* **90**, 491-500.

Farr, G.W., Furtak, K., Rowland, M.B., Ranson, N.A., Saibil, H.R., Kirchhausen, T. and Horwich, A.L. (2000) Multivalent binding of nonnative substrate proteins by the chaperonin GroEL. *Cell* **100**, 561-573.

Fayet, O., Ziegelhoffer, T. and Georgopoulos, C. (1989) The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J. Bacteriol.* **171,** 1379-1385.

Feierabend, J. (1975) Developmental studies on microbodies in wheat leaves. On the photo control of microbody development. *Planta* **123**, 63-77.

Feldman, D.E. and Frydman, J. (2000) Protein folding in vivo: the importance of molecular chaperones. *Curr. Opin. Stuctr. Biol.* 10, 26-33.

Geiser, J.R., Schott, E.J., Kingsbury, T.J., Cole, N.B., Totis, L.J., Bhattacharyya, G., He, L. and Hoyt, M.A. (1997) *Saccharyomyces cerevisiae* genes required in the absence of the CIN8-encoded spindle motor act in functionally diverse pathways. *Mol. Biol. Cell* **8**, 1035-1050.

Grellet, F., Raynal, M., Laudié, M., Cooke, R., Giraudat, J. and Delseny, M. (1993) cDNA encoding a putative 10-kilodalton chaperonin from *Arabidopsis thaliana*. *Plant Physiol.* **102**, 685.

Hedrick, J.L. and Smith, A.J. (1968) Size and Charge isomer separation and estimation

of molecular weights of proteins by disc gel electrophoresis. *Arch. Biochem. Biophys.* **123**, 155-164.

Hemmingsen, S.M., Woolford, C., van der Vies, S.M. Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W. and Ellis, R.J. (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333**, 330-334.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) A simple and general method for transferring genes into plants. *Science* **227**, 1229-1231.

Hodges, T.K. and Leonard, R.T. (1974) Cytochrome c oxidase. *Methods Enzymol.* **32,** 392-406.

Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, I. N. and Furtak, K. (1993) Folding in vivo of bacterial cytoplasmic proteins: Role of GroEL. *Cell* **74**, 909-917.

Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L. and Deisenhofer, J. (1996) The crystal structure of the GroES co-chaperonin at 2.8Å resolution. *Nature* **379**, 37-45

Hunt, J. F., van der Vies, S.,M., Henry, L. and Deisenhofer, J. (1997) Structural adaptations in the specialized bacteriophage T4 co-chaperonin Gp31 expand the size of the Anfinsen cage. *Cell* **90**, 361-371.

Kinoshita, T., Yamada, K., Hiraiwa, N., Kondo, M., Nishimura, M. and Hara-Nishimura, I. (1999) Vacuolar processing enzyme is up-regulated in the lytic vacuoles of vegetative tissues during senescence and under various stressed conditions. *Plant J.* **19**,43-53.

Koumoto, Y., Shimada, T., Kondo, M., Hara-Nishimura, I. and Nishimura, M. (2001) Chloroplasts have a novel Cpn10 in addition to Cpn20 as co-chaperonins in *Arabidopsis thalian*. 投稿中

Koumoto, Y., Shimada, T., Kondo, M., Takao, T., Shimonishi, Y., Hara-Nishimura, I. and Nishimura, M. (1999) Chloroplast Cpn20 forms a tetrameric structure in *Arabidopsis thalian*. *Plant J.* **17**, 467-477.

Koumoto, Y., Tsugeki, R., Shimada, T., Mori, H., Kondo, M., Hara-Nishimura, I. and Nishimura, M. (1996) Isolation and characterization of a cDNA encoding

mitochondrial chaperonin 10 from *Arabidopsis thaliana* by functional complementation of an *Escherichia coli groES* mutant. *Plant J.* **10**, 1119-1125.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Landry, S.J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C. and Gierasch, L.M. (1993) Characterization of a functionally important mobile domain of GroES. *Nature* **364**, 255-258.

Leroux, M.R., Fändrich, M., Klunker, D., Siegers, K., Lupas, A.N., Brown, J.R., Schiebel, E., Dobson, C.M. and Hartl, F.U. (1999) MtGimC, a novel archaeal chaperone related to the eukaryotic chaperonin cofactor GimC/prefoldin. *EMBO J.* **18**, 6730-3743.

Lubben, T.H., Gatenby, A.A., Donaldson, G.K., Lorimer, G.H. and Viitanen, P.V. (1990) Identification of a groES-like chaperonin in mitochondria that facilitates protein folding. *Proc. Natl. Acad. Sci. USA* **87**, 7683-7687.

Luck, H. (1965) Catalase. In Methods of Enzymatic Analysis. Edited by Bergmeyer, pp. 885-894. Academic Press, New York

Madueño, F., Napier, J.A. and Gray, J.C. (1993) Newly imported Rieske iron-sulfur protein associates with both Cpn60 and Hsp70 in the chloroplast stroma. *Plant Cell* **5**, 1865-1876.

Martel, R., Cloney, L.P., Pelcher, L.E. and Hemmingsen, S.M. (1990) Unique composition of plastid chaperonin-60:α andβ polypeptide-encoding genes are highly divergent. *Gene* **94** 181-187.

Mori, H., Takeda-Yoshikawa, Y., Hara-Nishimura, I. and Nishimura, M. (1991) Pumpkin malate synthase: Cloning and sequencing of the cDNA and Northern blot analysis. *Eur. J. Biochem.* **197**, 331-336.

Nishimura, M., Takeuchi, Y., De Bellis, L. and Hara-Nishimura, I. (1993) Leaf peroxisomes are directly transformed to glyoxysomes during senescence of pumpkin cotyledons. *Protoplasma* **175**, 131-137.

Nishio K, Hirohashi T, Nakai M. (1999) Chloroplast chaperonins: evidence for

heterogeneous assembly of alpha and beta Cpn60 polypeptides into a chaperonin oligomer. *Biochem Biophys Res Commun.* **266,** 584-587.

Prasad, T.K. and Stewart, C.R. (1992) cDNA clones encoding *Arabidopsis thaliana* and *Zea mays* mitochondrial chaperonin HSP60 and gene expression during seed germination and heat shock. *Plant Mol. Biol.* **18**, 873-885.

Rospert, S., Junne, T., Glick, B.S. and Schatz, G. (1993) Cloning and disruption of the gene encoding yeast mitochondrial chaperonin 10, the homolog of *E. coli* groES. *FEBS Lett.* **335**, 358-360.

Ryan, M.T., Hoogenraad, N.J. and Høj, P.B. (1994) Isolation of a clone specifying rat chaperonin 10, a stress-inducible mitochondrial matrix protein synthesised without a cleavable presequence. *FEBS Lett.* **337**, 152-156.

Schlicher, T. and Soll, J. (1996) Molecular chaperones are present in the thylakoid lumen of pea chloroplasts. *FEBS* **379**, 302-304.

Schmidt, M., Buchner, J., Todd, M.J., Lorimer, G.H. and Viitanen, P.V. (1994) On the groES in the chaperonin-assisted folding reaction. *J. Biol. Chem.* **269**, 10304-10311.

Soltys, B.J. and Gupta, R.S. (1999) Mitochondrial-matrix proteins at unexpected locations: are they exported? *Trends Biochem. Sci.* **24**.174-177.

Srere, P.A. (1969) Citrate synthase. *Methods Enzymol.* **13**, 3-11.

Tsugeki, R., Mori, H. and Nishimura, M. (1992) Purification, cDNA cloning and Northern-blot analysis of mitochondrial chaperonin 60 from pumpkin cotyledons. *Eur. J. Biochem.* **209**, 453-458.

Tsugeki, R. and Nishimura, M. (1993) Interaction of homologues of Hsp70 and Cpn60 with ferredoxin-$NADP^+$ reductase upon its import into chloroplasts. *FEBS Lett.* **320**, 198-202.

Vainberg, I.E., Lewis, S.A., Rommelaere, H., Ampe, C., Vandekerckhove, J,. Klein, H.L. and Cowan, N.J. (1998) Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell* **93**, 863-873.

Weissman, J.S., Hohl, C.M. Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H.R., Fenton, W.A. and Horwich, A.L. (1995) Mechanism of GroEL action: productive

release of polypeptide from a sequestered position under GroES. *Cell* **83**, 577-587.

Xu, Z., Horwich, A.L. and Sigler, P.B. (1997) The crystal structure of the asymmetric GroEL-GroES-$(ADP)_7$ chaperonin complex. *Nature* **388**, 741-750.

Zabaleta, E., Oropeza, A., Jimenez, B., Salerno, G., Crespi, M. and Herrera-Estrella, L. (1992) Isolation and characterization of genes encoding chaperonin 60β from *Arabidopsis thaliana*. *Gene* **111**, 175-181.

謝辞

本研究を遂行するにあたり、研究の機会を与えて下さり終始ご指導とご鞭撻を頂きました基礎生物学研究所の細胞機構研究部門の西村幹夫教授と、京都大学大学院植物学教室植物分子細胞生物学分野の西村いくこ教授に心から感謝いたします。本研究の大部分は基礎生物学研究所の細胞機構研究部門において行われました。ご助言、ご指導、激励を頂きました当研究部門の皆さまに感謝します。また、京都大学大学院植物学教室植物分子細胞生物学分野の皆さまにもご助言いただきました。お礼申しあげます。*groES* 変異体を分与して下さいました伊藤維昭先生（京都大学）感謝いたします。質量分析をして頂きました下西康嗣先生、高尾敏文先生（大阪大学）に感謝いたします。Functional cloning は槻木竜二先生（京都大学)、シロイヌナズナの cDNA ライブラリーの構築には森仁志先生（名古屋大学）のお力によるところが大きいことを申し添えます。免疫電顕は基礎生物学研究所の細胞機構研究部門の近藤真紀技官に、アミノ酸分析は大澤園子技官にご協力頂きました。ありがとうございました。タバコの形質転換をご指導頂いた木下哲博士、適切なご助言をして下さった山口勝司博士に感謝します。

報文目録

Koumoto, Y., Tsugeki, R., Shimada, T., Mori, H., Kondo, M., Hara-Nishimura, I. and Nishimura, M. (1996) Isolation and characterization of a cDNA encoding mitochondrial chaperonin 10 from *Arabidopsis thaliana* by functional complementation of an *Escherichia coli groES* mutant. *Plant J.* **10**, 1119-1125.

Koumoto, Y., Shimada, T., Kondo, M., Takao, T., Shimonishi, Y., Hara-Nishimura, I. and Nishimura, M. (1999) Chloroplast Cpn20 forms a tetrameric structure in *Arabidopsis thaliana*. *Plant J.* **17**, 467-477.

Koumoto, Y., Shimada, T., Kondo, M., Hara-Nishimura, I. and Nishimura, M. (2001) Chloroplasts have a novel Cpn10 in addition to Cpn20 as co-chaperonins in *Arabidopsis thalian*. 投稿中

The Plant Journal (1996) **10**(6), 1119–1125

Isolation and characterization of a cDNA encoding mitochondrial chaperonin 10 from *Arabidopsis thaliana* by functional complementation of an *Escherichia coli groES* mutant

Yasuko Koumoto, Ryuji Tsugeki[†], Tomoo Shimada, Hitoshi Mori[‡], Maki Kondo, Ikuko Hara-Nishimura and Mikio Nishimura*

Department of Cell Biology, National Institute for Basic Biology, Okazaki 444, Japan

Summary

Chaperonin (Cpn) is one of the molecular chaperones. Cpn10 is a co-factor of Cpn60, which regulates Cpn60-mediated protein folding. It is known that Cpn10 is located in mitochondria and chloroplasts in plant cells. The *Escherichia coli* homologue of Cpn10 is called GroES. A cDNA for the Cpn10 homologue was isolated from *Arabidopsis thaliana* by functional complementation of the *E. coli groES* mutant. The cDNA was 647 bp long and encoded a polypeptide of 98 amino acids. The deduced amino acid sequence showed approximately 50% identity to mammalian mitochondrial Cpn10s and 30% identity to GroES. A Northern blot analysis revealed that the mRNA for the Cpn10 homologue was expressed uniformly in various organs and was markedly induced by heat-shock treatment. The Cpn10 homologue was constitutively expressed in transgenic tobaccos. Immunogold and immunoblot analyses following the subcellular fractionation of leaves from transgenic tobaccos revealed that the Cpn10 homologue was localized in mitochondria and accumulated at a high level in transgenic tobaccos.

Introduction

Molecular chaperones mediate protein folding, assembly and protein translocation in eukaryotic and prokaryotic cells. One class of the molecular chaperones, the chaperonin (Cpn) family, consists of Cpn60 and Cpn10. Cpn60 and Cpn10 of *Escherichia coli*, namely, GroEL and GroES, respectively, have been characterized in detail. GroEL and GroES are heat-shock proteins that are also required for viability under normal conditions (Fayet *et al.*, 1989). It is thought that these molecular chaperones are induced by heat-denatured proteins and help them to regenerate. For many proteins, folding *in vivo* may not be a spontaneous process. Chaperonins particularly play a vital role in protein folding by their oligomeric structure; chaperonins construct protein-folding machinery (Gatenby and Viitanen, 1994). Cpn60 subunits form two stacked rings of the heptamer. This cylinder has a large cavity that captures incompletely folded proteins. Incompletely folded polypeptides expose their hydrophobic domains which are the interior regions of folded proteins. These polypeptides easily form aggregates under the high-concentration conditions in cells. These aggregations are prevented by binding of the polypeptides with Cpn60. Bound polypeptides are released from Cpn60 to yield biologically active proteins. This process is driven by the ATPase activity of Cpn60. As the Cpn60 ATPase activity is inhibited by binding with Cpn10, it is thought that Cpn10 is a regulator of this protein-folding machinery.

In plants, Cpn60 homologues are detected in mitochondria (Prasad and Stewart, 1992; Tsugeki *et al.*, 1992) and chloroplasts (Martel *et al.*, 1990). Mitochondrial Cpn60s show a high sequence similarity to bacterial Cpn60. A G–G–M motif is conserved at the carboxyl terminus of these Cpn60s. In chloroplasts, it has been reported that equal amounts of two distinct Cpn60 polypeptides (α and β) are present. Chloroplast Cpn10 has been reported to be comprised of two GroES-like domains following a transit peptide (Bertsch *et al.*, 1992). Grellet *et al.* (1993) reported the nucleotide sequence for a Cpn10 homologue of *Arabidopsis*, which is a single type with no transit peptide. However, its intracellular localization is not clear.

The bovine and rat mitochondrial Cpn10s were identified through experiments that showed that they could act as functional substitutes for GroES (Lubben *et al.*, 1990). It has been shown that the partially purified Cpn10 homologue in pea chloroplasts can also form a stable complex with GroEL and mediate protein folding with it (Bertsch *et al.*, 1992). In this paper, we tried functional cloning of cDNA for Cpn10 homologues using a library of cDNA from *Arabidopsis* and the *E. coli groES* mutant. As a result, a cDNA for the Cpn10 homologue was obtained. It is shown that the Cpn10 homologue could be transported into mitochondria when this cDNA was expressed in tobacco plant, indicating that the homologue is mitochondrial Cpn10 in *Arabidopsis*.

Received 5 April 1996; revised 30 August 1996; accepted 17 September 1996.
*For correspondence (fax +81 564 55 7505).
[†]Present address: Biotechnology Institute, 211 Wartik Laboratory, The Pennsylvania State University, University Park, PA 16802, USA.
[‡]Present address: School of Agriculture, Nagoya University, Chikusa, Nagoya 464–01, Japan.

Results

Functional cloning of cDNA for the Cpn10 homologue

The temperature-sensitive mutant strain of *E. coli*, YJ004 (KI471), can grow at 30°C but not at 42°C, because of a defect in *groES*, the Cpn10 homologue gene. An expression library was constructed with poly(A)$^+$ RNA from *Arabidopsis* using the expression vector pBluescript™ (pBS) and transformed into the temperature-sensitive *groES* mutant of *E. coli*. The transformed mutants harbouring the cDNA of Cpn10 homologues must complement the defective function of GroES to grow at 42°C. Each plasmid was prepared from the positive clones and was used to transform the *groES* mutant again. As shown in Figure 1, wild-type (1) and two positive clones (5 and 6) can grow at both 30°C and 42°C, whereas the *groES* mutant (2) and the mutants with pBS (3 and 4) can grow only at 30°C. The plasmid DNAs were isolated from the positive clones and their partial sequences were determined. A homology search revealed that cDNA inserts of the positive clones encoded the same Cpn10 homologue.

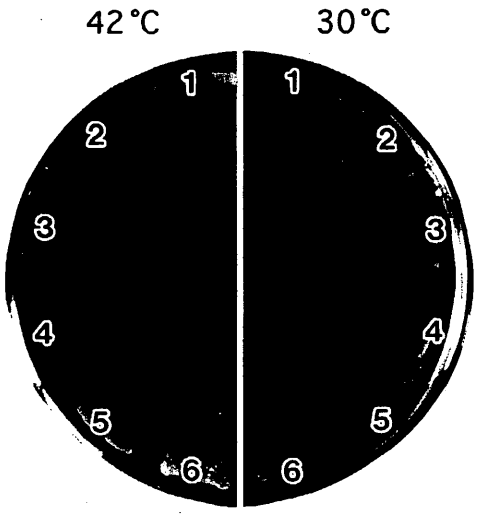


Figure 1. Functional complementation of temperature-sensitive *groES E. coli* mutant using the cDNA library from *Arabidopsis*.
Four kinds of *E. coli* cells were streaked on two plates. One plate was incubated at 30°C (right) and the other at 42°C (left). 1, Wild-type; 2, *groES* mutant; 3 and 4, mutants harbouring pBS; 5 and 6, mutants harbouring the cDNA of the Cpn10 homologue.

Sequence analysis of cDNA for the Cpn10 homologue

The complete nucleotide sequence of *Arabidopsis* Cpn10 homologue was determined for both strands. The nucleotide and deduced amino acid sequences are shown in

(a)

```
                                    gaagaattgagaa   13
ATGATGAAGCGTCTGATCCCAACGTTCAAC      43
 M  M  K  R  L  I  P  T  F  N       10
CGCATCTTGGTGCAGAGAGTCATCCAGCCC      73
- R  I  L  V  Q  R  V  I  Q  P      20
GCTAAAACCGAAAGCGGCATTCTCCTACCT     103
 A  K  T  E  S  G  I  L  L  P      30
GAGAAATCCTCCAAGCTGAACTCAGGCAAG     133
 E  K  S  S  K  L  N  S  G  K      40
GTGATAGCTGTTGGACCTGGATCAAGGGAT     163
 V  I  A  V  G  P  G  S  R  D      50
AAGGACGGGAAATTGATTCCGGTCTCTGTG     193
 K  D  G  K  L  I  P  V  S  V      60
AAGGAAGGCGACACTGTTCTTCTTCCAGAG     223
 K  E  G  D  T  V  L  L  P  E      70
TACGGTGGTACACAGGTCAAGCTCGGCGAG     253
 Y  G  G  T  Q  V  K  L  G  E      80
AACGAGTACCATCTCTTCCGGGACGAGGAT     283
 N  E  Y  H  L  F  R  D  E  D      90
GTTTTGGGAACTTTGCACGAGGATtgaaaa     313
 V  L  G  T  L  H  E  D  *         98
ggctaagcttgccaacttaaccacgagggt     343
tcatgttggtgtttgtgttatgaggagaag     373
tcatttataaattagtttatcttgaagatg     403
tggttggactttgttgtcgtttatcattga     433
atctacctttatgaacctgtctttgaattt     463
ttacaaatgggcatcaatcacatggataac     493
ccaagtgttgcatcttctcattttgtgctt     523
ttccgtaatctgtggatgcttttcgtttac     553
gtttaggcagcaacttcaatagccaatgta     583
gacgatagcttagttgactttgcttattca     613
aaaaaaaaaaaaaaaaaaaaaaaaaaaaa     643
aaaa                               647
```

(b)

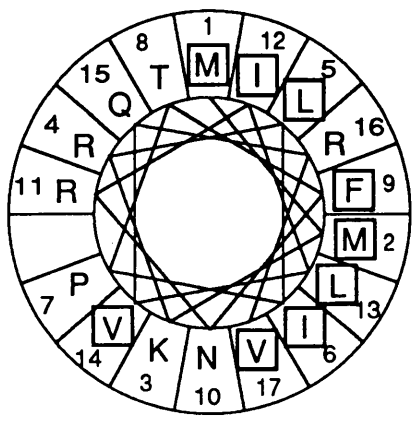


Figure 2. The nucleotide and deduced amino acid sequences of cDNA encoding the homologue of Cpn10 (a) and α-helical projection of the first 17 residues (b).
(a) Nucleotides are numbered from the 5′ to the 3′ region. The deduced amino acid sequence is numbered from the first methionine residue. Star indicates a termination codon.
(b) Helical wheel axial projection with 3.6 residues per turn. Hydrophobic residues are boxed.

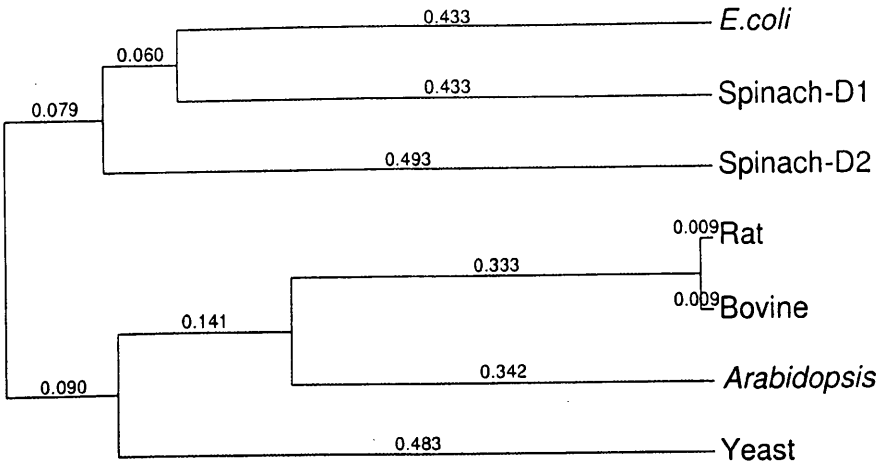


Figure 3. Relationships among the homologues of Cpn10.
The UPGMA program (GeneWorks; Intelligenetics, Mountain View, CA, USA) was used to calculate the relative lengths of the branches. The numbers indicate the validity of the branch point. The sequences represented the following studies: rat, Ryan *et al.* (1994); bovine, Pilkington and Walker (1993); spinach, Bertsch *et al.* (1992); yeast, Rospert *et al.* (1993); *E. coli*, Hemmingsen *et al.* (1988).

Figure 2(a). The cDNA is 647 bp long and encodes a polypeptide of 98 amino acids with the calculated molecular mass of 10 787 Da. The 3′ untranslated sequence consists of 340 nt, and a putative polyadenylation signal, TATAA, is located 235 nt upstream of the poly(A) tail. The nucleotide sequence is consistent with that reported by Grellet *et al.* (1993), but has an extra portion to the 3′ untranslated region. Figure 2(b) shows that the first 17 residues of the deduced amino acid sequence can form an amphiphilic α-helix. A phylogenic tree is established from amino acid sequences of Cpn10 homologues (Figure 3). As the spinach chloroplast Cpn10 has two GroES-like domains, the N-terminal portion (amino acids 9–104) and the C-terminal portion (amino acids 107–202) are designated D1 and D2, respectively, in accordance with an original paper (Bertsh *et al.*, 1992). The amino acid sequence shows a relatively high homology to Cpn10s of bovine and rat (49% and 48% identities, respectively), but low homology with those of yeast and *E. coli* (39% and 30%).

Expression of the His-tagged Cpn10 homologue in E. coli

cDNA for the Cpn10 homologue was expressed in *E. coli* in a His-tagged form. As shown in Figure 4, the His-tagged Cpn10 homologue was expressed in *E. coli* (lane T) and purified by column chromatography on Ni-NTA (lanes 3 and 4). Many other soluble proteins in *E. coli* did not bind to the Ni-NTA resin and appeared in the flow-through fraction (lane F). The purified His-tagged Cpn10 homologue showed slightly smaller mobility on SDS–PAGE, probably due to a His-tag and was used for preparation of the monospecific antibody. The His-tagged Cpn10 homologue behaved as an oligomer on gel-filtration column chromatography (data not shown). The molecular mass was estimated to be 70 000 Da from the elution profile, suggesting that the His-tagged Cpn10 homologue forms a heptameric structure.

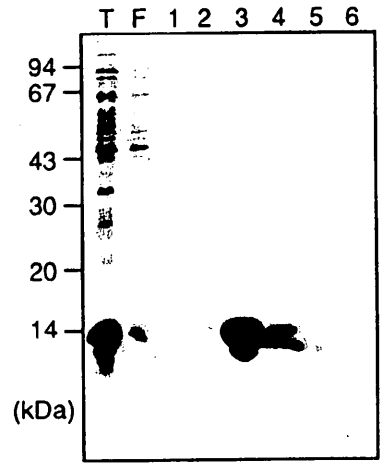


Figure 4. SDS–PAGE gel showing expression of the His-tagged Cpn10 homologue in *E. coli*.
The His-tagged Cpn10 homologue was expressed in *E. coli* and purified by column chromatography on Ni-NTA resin. The total proteins of extracts from *E. coli* that expressed the His-tagged Cpn10 homologue (lane T), the flow-through fraction (lane F) and fractions eluted with 0–0.5 M imidazole (lanes 1–6) were subjected to SDS–PAGE and subsequent staining with Coomassie Brilliant Blue.

Subcellular localization

To determine the localization of the Cpn10 homologue, subcellular fractionation of *Arabidopsis* leaves was performed. Organelles in leaf cells were separated on linear sucrose gradients and the distribution of each organelle was determined from those of marker enzyme activities. An immunoreactive band with an antiserum against the His-tagged Cpn10 homologue was detected at the peak density of 1.40 g cm^{-3} and co-fractionated with the mitochondrial marker, cytochrome c oxidase (data not shown). However, the activity of the chloroplastidal marker, triose phosphate isomerase, overlapped with that of cytochrome c oxidase. The chlorophyll profile revealed that there was no good separation between mitochondria and chloroplasts. Therefore, the fractionation experiments do not

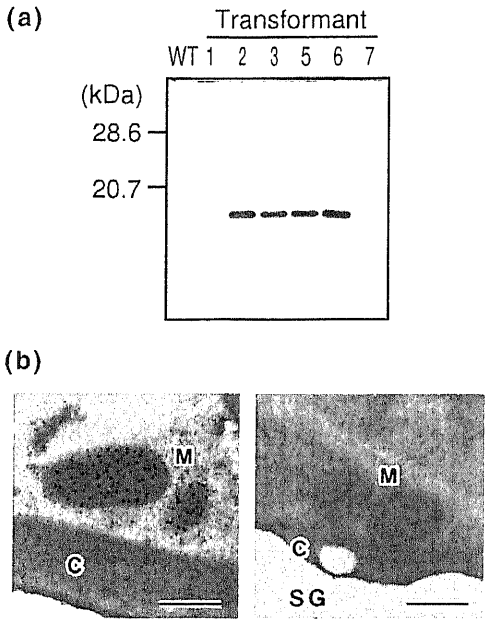


Figure 5. Immunoblotting (a) and immunogold labelling (b) of the transgenic tobacco that overproduces the Cpn10 homologue.
(a) The extracts of leaves of transformed and wild-type tobaccos were centrifuged and the protein concentrations of the supernatants were determined. Each proteins (10 μg) was subjected to SDS–PAGE and subsequent immunoblotting with an antiserum against the His-tagged Cpn10 homologue. Size markers are indicated at left.
(b) The transgenic tobaccos overproducing Cpn10 (left) and harbouring vectors (right) were grown in kanamycin- and hygromycin-containing media and transferred to soil. The sixth leaf from each flower bud was used for immunogold labelling with an antiserum against the His-tagged Cpn10 homologue. Bar, 0.5 μm; M, mitochondrion; C, chloroplast; SG, starch granule.

provide information on the subcellular localization of the Cpn10 homologue in *Arabidopsis*.

To clarify the localization of the Cpn10 homologue, transgenic tobacco plants harbouring the Cpn10 cDNA of *Arabidopsis* were prepared. The full length of the cDNA for the Cpn10 homologue was introduced into tobacco plants and expressed under regulation of the 35S promoter from cauliflower mosaic virus. All transformants were phenotypically normal. Since an antiserum against the Cpn10 homologue shows very little reactivity with endogenous Cpn10 in tobacco, only the introduced Cpn10 homologue was detected by immunoblotting. The Cpn10 homologue accumulated to a high level in most of the transformants (Figure 5a).

One of the transformants that overproduced the Cpn10 homologue (number 2 in Figure 5a) was subjected to subcellular fractionation and subsequent analysis as above. The Cpn10 homologue introduced into tobacco plant also co-fractionated with cytochrome c oxidase but not with triose phosphate isomerase (data not shown). This result suggested that the Cpn10 homologue was localized in mitochondria in the transgenic tobacco. Moreover, immunogold labelling of the leaves of the transformants confirmed

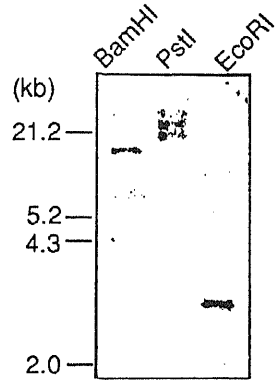


Figure 6. Genomic Southern blot analysis of the mitochondrial Cpn10 in *Arabidopsis*.
Restriction fragments were separated on an agarose-gel and subjected to subsequent Southern blotting using radio-labelled mit-Cpn10 cDNA as a probe. The restriction enzymes that were used are indicated above each lane. Size markers are indicated at the left.

that the expressed Cpn10 homologue was exclusively localized in mitochondria but not in chloroplasts and other organelles. An example of one such transformant (number 2) is shown in Figure 5(b). These results clearly show that the Cpn10 homologue is mitochondrial Cpn10 (mit-Cpn10) and is transported in mitochondria in tobacco plant.

Southern and Northern blot analyses

Genomic Southern blotting of *Arabidopsis* DNA was performed with radio-labelled mit-Cpn10 cDNA (Figure 6). After high stringent washing, a single band was observed in the *Eco*RI digest and two bands in the *Pst*I digest. In the *Bam*HI digest, there is one clear and one faint band. These results suggested that there is one or a few genes related to the mit-Cpn10 gene in the *Arabidopsis* genome. In the EST data base, there are a few cDNAs of the mit-Cpn10. We also found a cDNA sequence which is not identical but shows high sequence similarity to the mit-Cpn10 cDNA reported here. These observations suggest that there are at least two genes related to the mit-Cpn10 in *Arabidopsis*.

Figure 7(a) shows a Northern blot analysis of RNA extracted from roots, stems and leaves of *Arabidopsis* mature plants. The mRNA for mit-Cpn10 was expressed uniformly in various organs. Heat-responsive accumulation of the mRNA was investigated in *Arabidopsis*. The *Arabidopsis* plant was grown on soil at 22°C for 4 weeks and then transferred to 35°C for various times. RNA was extracted from the excised leaves. As shown in Figure 7(b), mRNA increased in the first hour after transfer and then high levels of mRNA continued to accumulate at least until 4 h after the temperature shift.

Discussion

We isolated a cDNA clone that functionally complemented the *E. coli groES* temperature-sensitive mutant from the

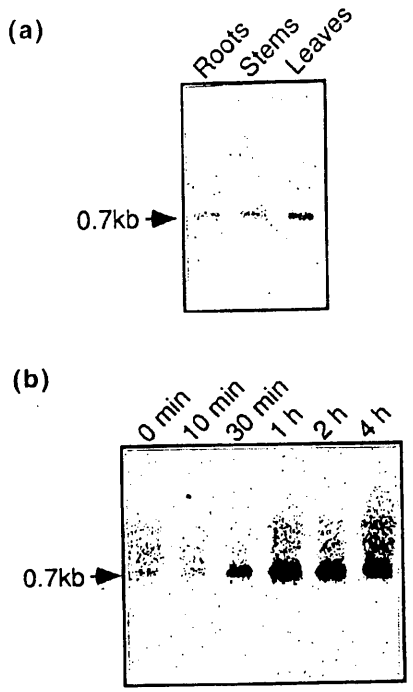


Figure 7. The organ specificity (a) and heat response (b) of mRNA for the mitochondrial Cpn10 in *Arabidopsis*.
(a) Total RNAs from roots, stems and leaves were subjected to Northern blotting using radio-labelled mit–Cpn10 cDNA as a probe. rRNA was used as an internal standard.
(b) *Arabidopsis* plants grown on soil were shifted from a normal growth temperature of 22°C up to 35°C (heat treatment). The total RNA from heat-treated leaves was subjected to Northern blotting as in (a). The heat-treatment time is shown above each lane.

Arabidopsis cDNA library. Sequence analysis revealed that this clone encoded a cDNA for a Cpn10 homologue. The amino acid sequence of the *Arabidopsis* Cpn10 homologue is more similar to that of mammalian mitochondrial Cpn10s than to that of the chloroplast Cpn10 of spinach. Furthermore, the amino acid sequence from Val66 to Val76 is completely identical with the partial sequence of potato mitochondrial Cpn10 (Burt and Leaver, 1994). From the analyses of subcellular fractionation and immunogold labelling of transgenic tobacco that expressed the Cpn10 homologue, it was confirmed that the Cpn10 homologue was localized in mitochondria. The deduced N-terminal amino acid sequence, from Met1 to Gln15, was expected to form an amphiphilic α-helix (Figure 2b), which has the potential to function as a target signal for mitochondria. This feature is also observed in mitochondrial Cpn10s of yeast (Rospert *et al.*, 1993) and a mammal (Ryan *et al.*, 1994).

A chloroplast Cpn10 homologue could not be isolated by the functional complementation of the *groES* mutant in this study. One possible reason for the inability to isolate it is the presence of a transit peptide in Chloroplast Cpn10, which may interfere with its function in *E. coli*. The spinach chloroplast Cpn10 homologue has been shown to have a transit peptide followed by two GroES-like domains fused together in tandem (Bertsch *et al.*, 1992). It has been shown that mature chloroplast Cpn10 of spinach is expressed in *E. coli* and complements the function of GroES (Baneyx *et al.*, 1995).

In order to prepare antibodies, the Cpn10 homologue was expressed in *E. coli*. Six histidine and some additional residues were fused to the N-terminus of the Cpn10 homologue to make it easier to purify. It was confirmed that this His-tagged Cpn10 homologue could form an oligomer on gel-filtration column chromatography. This result suggests that the His-tagged Cpn10 homologue represents its activity in *E. coli*. Antibodies prepared for this study were able to recognize *Arabidopsis* mitochondrial Cpn10, but not chloroplast Cpn10. It was reported that the antibodies against GroES cannot recognize mitochondrial Cpn10 in beef liver (Lubben *et al.*, 1990). In addition, our antibodies could not detect the mitochondrial Cpn10 homologue of tobacco. The former two cases show that Cpn10 homologues which have low identity to each other are hardly detected with same antibody.

In this study, the Cpn10 homologue gene was introduced in the sense orientation into tobacco plants and overexpression of the transgene was shown. From the results of subcellular fractionation and immunoelectron microscopy, it was confirmed that the Cpn10 homologue was also imported to mitochondria in transgenic tobacco leaves. These mitochondria containing the Cpn10 homologue were morphologically indistinguishable from mitochondria in control cells. The band of the introduced Cpn10 homologue on the immunoblot was 10 times as dense as that of wild-type *Arabidopsis*. The expression of the introduced Cpn10 homologue was constitutive and could not be induced by stress. This suggests that transgenic tobacco plants always have more mitochondrial Cpn10 than wild-type tobacco plants. In general, stress tolerance is acquired by mild stress pretreatment, because the pretreatment increases the quantity of molecular chaperones in cells (Lin *et al.*, 1984). As Cpn10 plays an important role in the regulation of Cpn60 which helps protein folding, overproduced Cpn10 might be effective in stress conditions that often cause refolding of proteins. We are now trying to determine the conditions where overproduced Cpn10 is most effective.

Experimental procedures

Construction of cDNA library

The *Arabidopsis* cDNA library in the *E. coli* expression vector pBluescript™ (Stratagene, La Jolla, CA, USA) was constructed by the modified vector-primer method of Mori *et al.* (1991). Poly(A)$^+$ RNA was prepared from 7-day-old dark-grown seedlings of *Arabidopsis thaliana*, ecotype Landsberg *erecta*.

Sequence analysis

Sequencing of cDNA encoding the homologue of Cpn10 was performed with a DNA sequencer (model 373A; Applied Biosystems Inc., Foster City, CA, USA) using M13 and Reverse fluorescent primers in accordance with the manufacturer's directions.

Expression and purification of the His-tagged Cpn10 homologue in Escherichia coli

The cDNA for the Cpn10 homologue was inserted into the expression vector pQE30 (QIAGEN, Chatsworth, CA, USA). Six histidine residues (His-tagged) located at the N-terminus of Cpn10 were derived from this vector. *Escherichia coli* were transformed with the constructs and the expression of proteins was induced with isopropyl-β-D-thiogalacto-pyranoside (IPTG). The extracts of sonicated *E. coli* cells were loaded on to a Ni-nitrilo-triacetic acid (Ni-NTA) column. The His-tagged Cpn10 homologue absorbed to the resin and eluted with a 30 ml gradient of 0–0.5 M imidazole after washing.

Plant transformation

The β-glucuronidase gene of a binary vector pBI121Hm was replaced by a *Xba*I–*Eco*RV cDNA fragment containing the Cpn10 gene. The pBI121Hm is a derivative of pBI121 and contains two drug resistance genes, the neomycin phosphotransferase II gene and the hygromycin phosphotransferase gene. The resulting construct was introduced into *Agrobacterium tumefaciens* (strain EHA101) and used to transform *Nicotiana tabacum* var. SR1 plants by the leaf disc method (Horsch *et al.*, 1985).

Subcellular fractionation and enzyme assays

Subcellular fractionations were performed as described by Tsugeki *et al.* (1992). In this case, the volumes of the gradient and the homogenate were 14 and 1.5 ml, respectively, and the centrifugation was performed at 85 500 ***g*** for 2.5 h with a Beckman SW28.1 rotor. Fractions were collected from the bottom of the gradient.

All assays of enzymatic activity were performed at room temperature. Enzymes were assayed as described in the cited references: cytochrome c oxidase (EC 1.9.3.1), Hodges and Leonard (1974); triose phosphate isomerase (EC 5.3.1.1), Feierabend (1975); catalase (EC 1.11.1.6), Luck (1965). Chlorophyll contents were determined by the method of Arnon (1949).

Electrophoresis and immunoblotting

SDS–PAGE (17.5% acrylamide) was performed by the method of Laemmli (1970). The separated proteins on the gels were electrophoretically blotted on to cellulose nitrate membranes. Immunochemical detection with monospecific antibody against the Cpn10 homologue was carried out with an ECL detection system (Amersham Japan, Tokyo, Japan).

Genomic Southern blotting

Total DNA was isolated from *Arabidopsis* by the cetyltrimethyl ammonium-bromide method. Total DNA (2 μg) was digested with restriction enzymes. The digests were fractionated on a 0.8% agarose gel and transferred to a nylon membrane under vacuum with 0.25 M NaOH and 1.5 M NaCl. Probes were prepared by the random-primer method with Cpn10 cDNA inserts as templates. Hybridization was performed as described by Church and Gilbert (1984). After hybridization, the membrane was sequentially washed in 2× SSC and 0.1% SDS at 42°C for 30 min and in 0.1× SSC and 0.1% SDS at 65°C for 30 min. The washed membrane was exposed to an imaging plate and the radioactivities of the signals were detected with the Analyzer system (Fuji Film, Tokyo, Japan).

Northern blotting

Total RNA was purified by LiCl precipitation following the extraction of RNA with ISOGEN (Nippongene, Tokyo, Japan). Total RNA (10 μg) was electrophoresed on a 1% agarose/formaldehyde gel and transferred to a nylon membrane under vacuum with 1 M ammonium acetate. Hybridization was performed as described above. After hybridization, the membrane was washed in 0.1× SSC and 0.1% SDS at 65°C for 3 h and the radioactivities of the signals were detected as described above.

Immunogold localization

Tobacco leaves were fixed, dehydrated and embedded in LR White resin (London Resin Co., Basingstoke, UK) as described previously (Nishimura *et al.*, 1993). Immunogold procedures were essentially the same as those described by Kato *et al.* (1996), except for the use of antiserum against the His-tagged Cpn10 homologue (diluted 500-fold) and 30-fold diluted protein A-gold (Amersham Japan, Tokyo, Japan). The sections were examined with a transmission electron microscope (1200EX; JEOL, Tokyo, Japan) at 80 kV.

Acknowledgements

We wish to thank Dr Koreaki Ito, Kyoto University, who kindly provided us with the *E. coli groES* mutant. This work was supported by Grants-in-Aid for Scientific Research (B) (No. 07456053) and on Priority Areas (Nos. 04273101, 04273102) from the Ministry of Education, Science and Culture, Japan and by a grant from the Nissan Science Foundation (Tokyo, Japan).

References

Arnon, D.I. (1949) Copper enzymes in chloroplasts. Polyphenol oxidase in *Beta vulgaris*. *Plant Physiol.* **24**, 1–15.

Baneyx, F., Bertsch, U., Kalbach, C.E., van der Vies, S.M., Soll, J. and Gatenby, A.A. (1995) Spinach chloroplast cpn21 co-chaperonin possesses two functional domains fused together in a toroidal structure and exhibits nucleotide-dependent binding to plastid chaperonin 60. *J. Biol. Chem.* **270**, 10 695–10 702.

Bertsch, U., Soll, J., Seetharam, R. and Viitanen, P.V. (1992) Identification, characterization, and DNA sequence of a functional 'double' groES-like chaperonin from chloroplasts of higher plants. *Proc. Natl Acad. Sci. USA*, **89**, 8696–8700.

Burt, W.J.E and Leaver, C.J. (1994) Identification of a chaperonin-10 homologue in plant mitochondria. *FEBS Lett.* **339**, 139–141.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci. USA*, **81**, 1991–1995.

Fayet, O., Ziegelhoffer, T. and Georgopoulos, C. (1989) The *groES* and *groEL* heat shock gene products of *Escherichia coli* are

essential for bacterial growth at all temperatures. *J. Bacteriol.* **171**, 1379–1385.

Feierabend, J. (1975) Developmental studies on microbodies in wheat leaves. On the photo control of microbody development. *Planta,* **123**, 63–77.

Gatenby, A.A. and Viitanen, P.V. (1994) Structural and functional aspects of chaperonin-mediated protein folding. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **45**, 469–491.

Grellet, F., Raynal, M., Laudié, M., Cooke, R., Giraudat, J. and Delseny, M. (1993) cDNA encoding a putative 10-kilodalton chaperonin from *Arabidopsis thaliana. Plant Physiol.* **102**, 685.

Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W. and Ellis, R.J. (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature,* **333**, 330–334.

Hodges, T.K. and Leonard, R.T. (1974) Cytochrome c oxidase. *Methods Enzymol.* **32**, 392–406.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) A simple and general method for transferring genes into plants. *Science,* **227**, 1229–1231.

Kato, A., Hayashi, M., Kondo, M. and Nishimura, M. (1996) Targeting and processing of a chimeric protein with the amino-terminal presequence of the precursor to glyoxysomal citrate synthase. *Plant Cell,* **8**, 1601–1611.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* **227**, 680–685.

Lin, C.-Y., Roberts, J.K. and Key, J.L. (1984) Acquisition of thermotolerance in soybean seedlings. *Plant Physiol.* **74**, 152–160.

Lubben, T.H., Gatenby, A.A., Donaldson, G.K., Lorimer, G.H. and Viitanen, P.V. (1990) Identification of a groES-like chaperonin in mitochondria that facilitates protein folding. *Proc. Natl Acad. Sci. USA,* **87**, 7683–7687.

Luck, H. (1965) Catalase. In *Methods of Enzymatic Analysis.* (Bergmeyer, H.U., ed.). New York: Academic Press, pp. 885–894.

Martel, R., Cloney, L.P., Pelcher, L.E. and Hemmingsen, S.M. (1990) Unique composition of plastid chaperonin-60: α and β polypeptide-encoding genes are highly divergent. *Gene,* **94**, 181–187.

Mori, H., Takeda-Yoshikawa, Y., Hara-Nishimura, I. and Nishimura, M. (1991) Pumpkin malate synthase: Cloning and sequencing of the cDNA and Northern blot analysis. *Eur. J. Biochem.* **197**, 331–336.

Nishimura, M., Takeuchi, Y., De Bellis, L. and Hara-Nishimura, I. (1993) Leaf peroxisomes are directly transformed to glyoxysomes during senescence of pumpkin cotyledons. *Protoplasma,* **175**, 131–137.

Pilkington, S.J. and Walker, J.E. (1993) Complementary DNA sequence of bovine cpn10 (HSP10), a chaperone protein from mitochondria. *DNA Sequencing,* **3**, 291–295.

Prasad, T.K. and Stewart, C.R. (1992) cDNA clones encoding *Arabidopsis thaliana* and *Zea mays* mitochondrial chaperonin HSP60 and gene expression during seed germination and heat shock. *Plant Mol. Biol.* **18**, 873–885.

Rospert, S., Junne, T., Glick, B.S. and Schatz, G. (1993) Cloning and disruption of the gene encoding yeast mitochondrial chaperonin 10, the homolog of *E. coli* groES. *FEBS Lett.* **335**, 358–360.

Ryan, M., Hoogenraad, N.J. and Høj, P.B. (1994) Isolation of a cDNA clone specifying rat chaperonin 10, a stress-inducible mitochondrial matrix protein synthesised without a cleavable presequence. *FEBS Lett.* **337**, 152–156.

Tsugeki, R., Mori, H. and Nishimura, M. (1992) Purification, cDNA cloning and Northern-blot analysis of mitochondrial chaperonin 60 from pumpkin cotyledons. *Eur. J. Biochem.* **209**, 453–458.

EMBL, GenBank and DDBJ data library accession number D88314 for the nucleotide sequence data reported in this paper.

The Plant Journal (1999) **17**(5), 467–477

Chloroplast Cpn20 forms a tetrameric structure in *Arabidopsis thaliana*

Yasuko Koumoto[1], Tomoo Shimada[1], Maki Kondo[1], Toshifumi Takao[2], Yasutsugu Shimonishi[2], Ikuko Hara-Nishimura[1] and Mikio Nishimura[1,*]

[1]*Department of Cell Biology, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji, Okazaki, 444–8585, Japan, and*

[2]*Instisute for Protein Research, Osaka University, Suita, 565–0871, Japan*

Summary

Chloroplast chaperonin 20 (Cpn20) in higher plants is a functional homologue of the *Escherichia coli* GroES, which is a critical regulator of chaperonin-mediated protein folding. The cDNA for a Cpn20 homologue of *Arabidopsis thaliana* was isolated. It was 958 bp long, encoding a protein of 253 amino acids. The protein was composed of an N-terminal chloroplast transit peptide, and the predicted mature region comprised two distinct GroES domains that showed 42% amino acid identity to each other. The isolated cDNA was constitutively expressed in transgenic tobacco. Immunogold labelling showed that Cpn20 is accumulated in chloroplasts of transgenic tobacco. A Northern blot analysis revealed that mRNA for the chloroplast Cpn20 is abundant in leaves and is increased by heat treatment. To examine the oligomeric structure of Cpn20, a histidine-tagged construct lacking the transit peptide was expressed in *E. coli* and purified by affinity chromatography. Gel-filtration and cross-linking analyses showed that the expressed products formed a tetramer. The expressed products could substitute for GroES to assist the refolding of citrate synthase under non-permissive conditions. The analysis on the subunit stoichiometry of the GroEL–Cpn20 complex also revealed that the functional complex is composed of a GroEL tetradecamer and a Cpn20 tetramer.

Introduction

Chaperonins are proteins that play a vital role in protein folding in eukaryotic and prokaryotic cells. The best studied chaperonin system is the GroEL–GroES complex from *Escherichia coli*. These proteins are stress-inducible and referred to as heat shock proteins. GroEL is composed of two heptameric rings, each consisting of seven 57 kDa subunits. Two stacked rings of the GroEL heptamer form a central cavity that captures incompletely folded proteins. The co-chaperonin GroES forms an asymmetric complex with the GroEL in which a single heptameric ring of GroES binds to the end of the GroEL cylinder. When both a polypeptide and GroES bind to the same end of GroEL, the ATP-dependent release of the polypeptide promotes the folding to its native form (Weissman *et al.*, 1995). GroES has been shown to increase the co-operativity of the ATPase activity of GroEL (Todd *et al.*, 1994). The crystal structures of the GroEL tetradecamer (Braig *et al.*, 1994), the GroES heptamer (Hunt *et al.*, 1996) and the GroEL–GroES complex (Xu *et al.*, 1997) are also documented in the literature. The structure of the complex shows that the mutationally identified residues responsible for peptide binding are removed from the surface of a cavity by binding of GroES, so that the central cavity becomes polar to favour the native protein. It has also been shown that GroES binding causes the steric block to ADP release (Xu *et al.*, 1997). Thus, the molecular mechanism of GroEL–GroES-assisted protein folding becomes increasingly elucidated.

In higher plants, chaperonin 60s (Cpn60s), which are closely related GroEL homologues, are found in mitochondria and chloroplasts. Chloroplast Cpn60 was originally identified as the ribulose bisphosphate carboxylase (Rubisco) subunit binding protein. Cpn60s can form stable complexes in an ATP-dependent manner with proteins imported into isolated chloroplasts, e.g. the Rieske iron-sulfur protein or ferredoxin-NADP$^+$ reductase (Madueño *et al.*, 1993; Tsugeki and Nishimura, 1993). In contrast to the single-type Cpn60 polypeptide in mitochondria, chloroplasts contain equal amounts of two distinct Cpn60 polypeptides (α and β). Their cDNAs have been cloned and sequence analyses indicated that Cpn60α and Cpn60β are very divergent (Hemmingsen *et al.*, 1988; Martel *et al.*, 1990; Zabaleta *et al.*, 1992). It is not known whether the α and β polypeptides form homo- or hetero-tetradecamers. Cpn60α and Cpn60β cDNAs of *Brassica napus* have been expressed separately or in combination in *E. coli* (Clony *et al.*, 1992). Cpn60β is efficiently assembled into a tetradecamer in *E. coli* both in the presence and absence of Cpn60α. In contrast, with Cpn60α such efficient assembly occurs only when Cpn60β is co-expressed. Thus, it is suggested that certain functions of Cpn60α and Cpn60β are not interchangeable.

Cpn10s, which are GroES homologues, have also been detected in both mitochondria and chloroplasts of higher plants. The cDNA for mitochondrial Cpn10 of *Arabidopsis thaliana* has been cloned by complementation of a GroES-

Received 10 August 1998; revised 14 December 1998; accepted 14 December 1998.

*For correspondence (fax + 81 564 55 7505; e-mail mikosome@nibb.ac.jp).

deficient *E. coli* (Koumoto *et al.*, 1996). Mitochondrial Cpn10 has no cleavable presequence. Instead, the N-terminal amino acid sequence forms an amphiphilic α-helix that has the potential to function as a targeting signal for mitochondria. This feature is also observed in the homologues of yeast (Rospert *et al.*, 1993) and rat (Ryan *et al.*, 1994). The chloroplast Cpn20, a functional homologue of Cpn10, has been partially purified from pea, and the cDNA of the spinach homologue has been cloned (Bertsch *et al.*, 1992). Interestingly, the chloroplast Cpn20 comprises two GroES-like domains following a transit peptide and has twice the size of GroES and mitochondrial Cpn10s; hence its designatation as Cpn20s (Hartl, 1996). Recently, some aspects of the roles of the different domains in chloroplast Cpn20 have been characterized (Baneyx *et al.*, 1995; Bertsch and Soll, 1995). The cDNAs carrying (i) the mature Cpn20, i.e. the complete sequence with tandem GroES-like domains; (ii) the N-terminal GroES domain by itself; and (iii) the C-terminal GroES-domain by itself, were expressed in GroES-deficient strain of *E. coli*. The double-domain Cpn20 was shown to be functional in *E. coli* and each of the N- and C-terminal GroES domains were also shown to be functional, although they were less effective than the double-domain Cpn20. However, *in vitro* experiments with GroEL, namely experiments involving assisted protein folding and ATPase inhibition, showed that both the N- and C-terminal GroES domains by themselves are defective (Bertsch and Soll, 1995).

Here we report the cDNA sequence for the Cpn20 homologue of *A. thaliana*. As expected this clone encoding the double-domain Cpn20 and its products was transported to chloroplasts in leaves of transgenic tobacco plants. As two kinds of chloroplast Cpn60, namely α and β, form a tetradecamer, the oligomeric structure of Cpn20 with the tandem GroES-like domain was of interest. Although the function of Cpn20 in the assembly of the protein has already been characterized, the structural information of its oligomer is as yet unknown. Cpn20 cDNA was expressed in *E. coli* and the oligomeric structure of the expressed products was studied.

Results and discussion

Cloning of a cDNA for the Cpn20 homologue

As described in the Experimental procedures, the *Arabidopsis* cDNA library was screened for the Cpn20 homologue using DIG-labelled 185 bp PCR products. The isolated cDNA was 958 bp long encoding a protein of 253 amino acids. A comparison of the amino acid sequence with that of a spinach chloroplast Cpn20 suggested that the N-terminal region of 51 amino acids is a chloroplast transit peptide. As in the case of spinach Cpn20, *Arabidopsis* Cpn20 was composed of two GroES-like domains fused in tandem. The predicted mature region of the *Arabidopsis* Cpn20 homologue was 61% identical to that of spinach Cpn20. The amino acid sequences of the N-terminal domain (residues 60–154) and the C-terminal domain (residues 158–253) showed 30% and 37% identities to a mitochondrial homologue of *Arabidopsis*. The two domains of the *Arabidopsis* Cpn20 homologue showed 42% amino acid identity to each other.

The deduced amino acid sequence of the mature regions of a mitochondrial Cpn60 (Cole *et al.*, 1994) and of two chloroplast Cpn60s (Cpn60α and Cpn60β) (Martel *et al.*, 1990) from *B. napus* were compared with each other. The identities between the mitochondrial Cpn60 and Cpn60α, the mitochondrial Cpn60 and Cpn60β, and Cpn60α and Cpn60β were 41%, 42% and 50%, respectively. This suggests that the degrees of divergence between the mitochondrial Cpn10 and the two domains of Cpn20 are similar to the degrees of divergence between the three isoforms of Cpn60.

Localization of Arabidopsis *Cpn20*

We have previously revealed by immunogold labelling that the *Arabidopsis* Cpn10 expressed in transgenic tobaccos is localized in mitochondria (Koumoto *et al.*, 1996). As tobacco endogenous chloroplast Cpn20 was hardly detectable by immunogold labelling, transgenic tobacco plants that overproduced Cpn20 were prepared to confirm the localization of the Cpn20. The full length of the cDNA for a chloroplast Cpn20 was introduced into tobacco plants and expressed under regulation of the 35S promoter from cauliflower mosaic virus. An immunoblot analysis of six transformants and the tobacco carrying the vector is shown in Figure 1(a). Two immunoreactive bands of closely related size were detected in a control plant (lane labelled 'V'), which indicated that the tobacco plant might have at least two isoforms of Cpn20. Some transformants accumulated the *Arabidopsis* Cpn20 in larger amounts (Figure 1a; lanes 1–3, 5–1, 6–1, 8–1 and 8–2).

In one of the transformants that overproduced the *Arabidopsis* Cpn20, the gold particles for a chloroplast Cpn20 were exclusively restricted to chloroplasts and were not detected in other organelles such as mitochondria or microbodies (Figure 1b). Since the gold particles were not specifically detected in any other organelle in the transformant that carried the vector (Figure 1c), it was concluded that the *Arabidopsis* Cpn20 introduced into tobacco plants accumulated in the chloroplasts. No differences were observed morphologically between the transformants and non-transformants. Recently, it has been shown that chloroplasts contain two types of co-chaperonin with different suborganellar localization: one is localized in the stroma and the other is localized in the thylakoid lumen (Schlicher and Soll, 1996). Schlicher and Soll (1996)

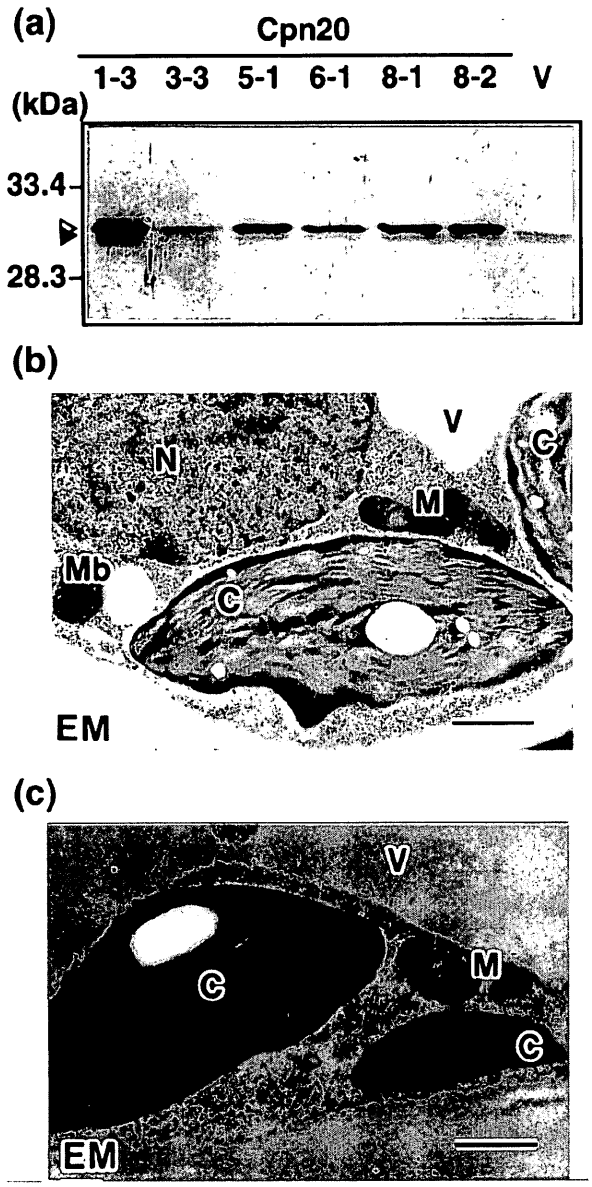


Figure 1. Immunoblotting (a) and immunogold labelling (b and c) of the transgenic tobacco plants.
(a) The extracts of leaves of transgenic tobacco plants were centrifuged and the protein concentrations of the supernatants were determined. Each protein fraction (10 μg) was subjected to SDS–PAGE and subsequent immunoblotting with an antiserum against the His-tagged Cpn20. The numbers denote the transgenic lines and 'V' indicates the transgenic tobacco carrying the vector only. Size markers are indicated at the left. The positions of migration of endogenous Cpn20 (closed triangle) and introduced *Arabidopsis* Cpn20 (open triangle) are also shown at the left.
(b and c) The transgenic tobacco plants were grown in kanamycin- and hygromycin-containing medium. One leaf was used for immunogold labelling with an antiserum against the His-tagged Cpn20. Immunogold labelling of the transgenic tobacco plant that overproduces chloroplast Cpn20 is shown in (b) and that of the control plant is shown in (c).
Bar, 1 μm: C, chloroplast; M, mitochondrion; Mb, microbody; N, nucleolus; V, vacuole; EM, extracellular matrix.

reported that the stromal co-chaperonin is a double-domain Cpn20 and the thylakoid lumenal one is of 'normal' size Cpn10. Immunogold labelling of the *Arabidopsis* Cpn20 revealed that the chloroplast Cpn20 is localized in the stroma, which matches the finding that the chloroplast co-chaperonin with a double-domain is a stromal protein.

Southern and Northern analyses

To investigate the number of related genes in *Arabidopsis*, genomic Southern blotting of *Arabidopsis* DNA was performed with a radiolabelled chloroplast Cpn20 cDNA (Figure 2a). After high stringent washing, a few bands were observed. Each of the digests, except for the *Pst*I digest, gave both strong and faint bands. These results suggest that there is one or a few genes in the *Arabidopsis* genome that is related to the Cpn20 gene. However, in the EST database we were unable to find a cDNA sequence that was very similar but not identical to the Cpn20 sequence.

The patterns of expression of the Cpn20 gene in plants were also examined. Figure 2(b) shows Northern blotting of RNA extracted from roots, stems and leaves of mature *Arabidopsis* plants. The mRNA for chloroplast Cpn20 was abundant in leaves and was also present in roots and stems. These results indicate that Cpn20 exists not only in chloroplasts but also in etioplasts. The effect of heat treatment on the accumulation of Cpn20 mRNA is shown in Figure 2(c). *Arabidopsis* plants that had been grown on soil at 22°C were transferred to 35°C for various lengths of time and then RNA was extracted from the excised leaves. Heat treatment resulted in a rapid increase in the amounts of Cpn20 mRNA accumulation, which lasted for at least 4 h. Although the mRNA content of the mitochondrial Cpn10 also increased by heat treatment, the level of Cpn10 mRNA remained constant from 1 to 4 h after the beginning of the temperature shift (Koumoto *et al.*, 1996).

Analysis of the oligomeric structure of the His-tagged chloroplast Cpn20

It is known that Cpn10 (GroES) functions as a heptamer. However, the chloroplast Cpn20 has two domains, so it is unlikely that it forms the same oligomeric structure by itself. For examination of the oligomeric structure of *Arabidopsis* Cpn20, the mature region of Cpn20 was expressed in *E. coli*. The transit peptide and the seven N-terminal amino acids of the predicted mature region (ASVVAPK) were truncated and a segment containing the histidine tag (MRGSHHHHHHG) was fused to the N-terminus. The expressed products had the residues PLXD (residues 64–67), which were highly conserved at the N-terminal region and were almost the same length as the predicted mature Cpn20. His-tagged Cpn20 was expressed in a soluble form and purified by Ni-NTA column chromatography. This

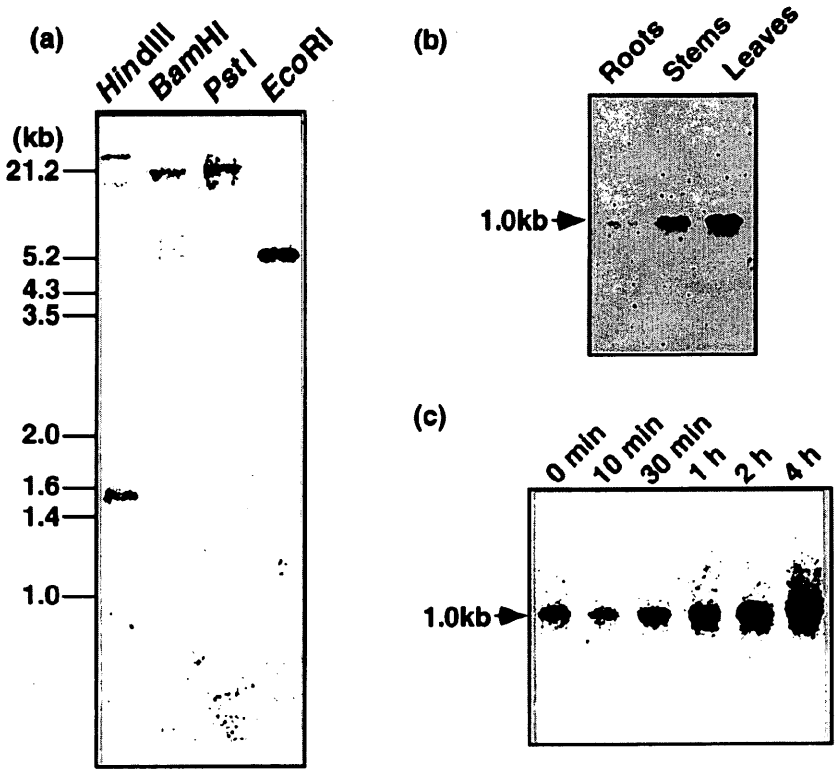


Figure 2. Genomic Southern blot analysis (a) and Northern blot analysis (b, c).
(a) *Hin*dIII, *Bam*HI, *Pst*I and *Eco*RI restriction fragments were separated on an agarose gel and subjected to subsequent Southern blotting probed with a radiolabelled chloroplast Cpn20 cDNA.
(b) The total RNA from roots, stems and leaves was subjected to Northern blotting probed with a radiolabelled chloroplast Cpn20 cDNA.
(c) *Arabidopsis* plants grown on soil were transferred from a normal growth temperature of 22°C to 35°C (heat treatment). The total RNA from heat-treated leaves was subjected to Northern blotting as in (a). The heat-treatment time is shown above each lane.

eluent mainly contained His-tagged Cpn20, as judged from the Coomassie Brilliant Blue (CBB) staining pattern (Figure 3a, insert), and was used for further analyses.

To analyse the oligomeric structure of the His-tagged Cpn20, gel-filtration column chromatography was performed. It was expected that the Cpn20 oligomer might not have a globular shape and thus would behave differently compared with the standard proteins. Therefore, the calculated molecular masses discussed below were simply used for comparison. The His-tagged chloroplast Cpn20 eluted from the Ni-NTA column was loaded and the absorbance at 280 nm was monitored (Figure 3a, top). A common peak in the two traces appearing at ≈ 34 min might be due to soluble proteins from *E. coli*, a contaminant during the fractionation on the Ni-NTA column. The main peak appeared at ≈ 25 min and the second peak appeared at ≈ 30 min, corresponding to molecular masses of 75 kDa and 16 kDa, respectively. As the protein with the molecular mass of 16 kDa seemed to be a monomer, the oligomeric form of the His-tagged Cpn20 was calculated to be a tetramer or a pentamer. For further confirmation, the same analysis was performed with a longer column. As shown in Figure 3(b), two peaks corresponding to molecular masses of 86 kDa and 18 kDa were observed. This result also indicates that the oligomeric form of the His-tagged Cpn20 is a tetramer or a pentamer. The His-tagged mitochondrial Cpn10 that had been prepared previously was also analysed as a control. Two peaks corresponding to molecular masses of 84 kDa and 12 kDa were observed (Figure 3a, top). This indicates that His-tagged Cpn10 forms a heptameric structure identical to that of GroES (Hunt *et al.*, 1996).

For determination of the oligomerization state, a cross-linking method was employed. To remove imidazole, His-tagged Cpn20 and Cpn10 were subjected to a desalting column. The desalted fraction was cross-linked with various concentrations of cross-linker (DSP) in 50 mM phosphate buffer, pH 7.2, containing 0.1 M NaCl, and analysed by SDS–PAGE (Figure 4a). DSP preferentially reacts with primary amines, including a terminal α-amino group and the ε-amino group of lysine. At high concentrations of DSP the larger oligomers of His-tagged Cpn20 and Cpn10 were observed. The largest oligomer seemed to be a tetramer in the case of Cpn20 and a heptamer for Cpn10. The relative mobility of a Cpn10 heptamer was slightly slower than that of a Cpn20 tetramer, which is consistent with the results

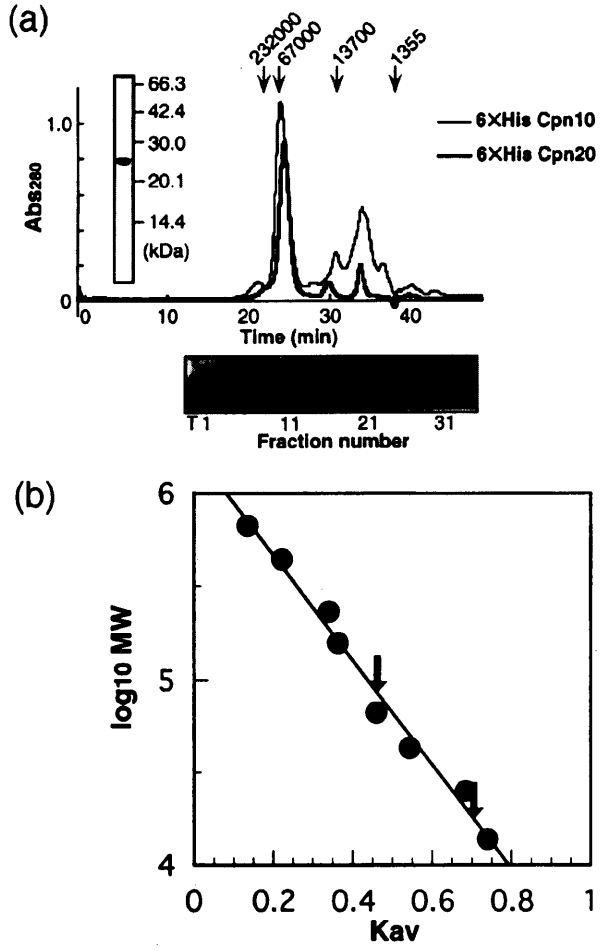


Figure 3. Molecular masses of oligomers of co-chaperonin homologues.
(a) His-tagged Cpn20 and Cpn10 were analysed by gel-filtration column chromatography (top). SDS–PAGE of the preparation used for the experiment was performed and stained with CBB. The CBB stained gel is inserted. The bold line indicates the absorbance monitored at 280 nm of His-tagged Cpn20 and a normal line indicates His-tagged Cpn10. Peak positions of standards (see the Experimental procedures) with their molecular masses are shown by the arrows. A common peak in the two lines appearing at ≈ 34 min might be due to a soluble protein from *E. coli*, a contaminant from the fractionation on the Ni-NTA column. After separation of the soluble proteins of tobacco chloroplasts, fractions with odd numbers were subjected to SDS–PAGE and subsequent immunoblotting with an antiserum against the His-tagged Cpn20 (bottom). The immunoblot of the total soluble proteins is shown at the left (T).
(b) The calibration curve using molecular mass standards (circle, see the Experimental procedures) on a HiLaod Superdex 200 pg column is shown. The calibration curve ($y = -2.81x + 6.23$) was calculated by the method of least squares. The Kav values of His-tagged Cpn20 (0.461 and 0.705) are shown by arrows.

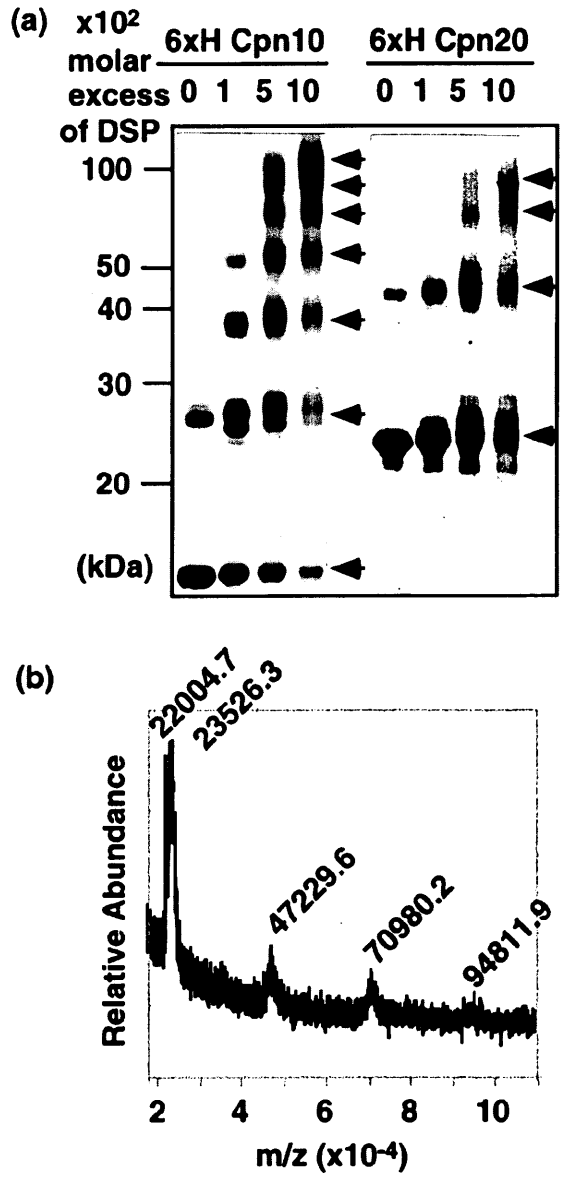


Figure 4. Oligomer structures of co-chaperonin homologues.
(a) His-tagged Cpn20 (right) and Cpn10 (left) were incubated with various concentrations of DSP omitting the reducing agent, 2ME, and analysed by SDS–PAGE followed by CBB staining. Arrowheads show positions of monomer and oligomer bands. Size markers are given at the left.
(b) The His-tagged Cpn20 oligomers were cross-linked at a 1000-fold molar excess of DSP and analysed by mass spectrometry. Numbers represent the molecular mass of the peak signals.

obtained with gel-filtration column chromatography. Even at a DSP concentration corresponding to a 1000-fold molar excess over the protein concentration smaller oligomers were still present. Because higher concentrations of DSP caused protein aggregation, we could not obtain complete cross-linkage. It has been reported that an equilibrium exists between a tetramer and heptamer of the *M. tubercu-losis* Cpn10 protein that can be modulated by the addition of divalent cations (Fossati *et al.*, 1995). To analyse such an effect on the oligomerization, we performed cross-linking analysis in the presence of magnesium ions. The data clearly showed that His-tagged Cpn20 formed a tetramer in the presence of 10 mM magnesium ions with the same efficiency as in the absence of magnesium ions (data not shown).

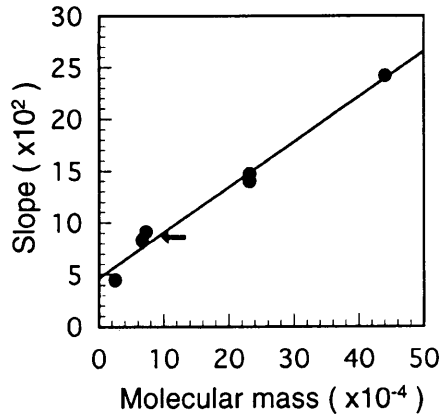


Figure 5. The slope–molecular mass relationship of standard proteins and the molecular mass determination of His-tagged Cpn20.
The slopes were calculated from the plots of the log of protein mobilities versus acrylamide gel concentration. Standard marker proteins (circle, see the Experimental procedures) were used for calibration. The calibration curve ($y = 0.438x + 4.66$) was calculated by the method of least squares. The slope of His-tagged Cpn20 (8.6) is shown by arrow.

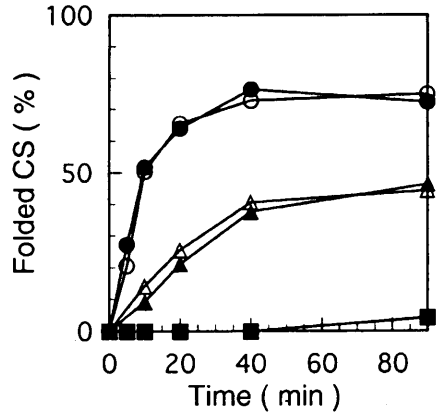


Figure 6. Time–course of folding of citrate synthase in the presence of His-tagged co-chaperonins.
Unfolded citrate synthase (CS) was diluted to a concentration of 150 nM with a solution containing $GroEL_{14}$ at 0°C. The temperature was immediately adjusted to 35°C and the following additions were made: square, 2 mM ATP; open triangle, 2 mM ATP and 300 nM His-tagged $Cpn20_4$; closed triangle, 2 mM ATP and 600 nM His-tagged $Cpn20_4$; open circle, 2 mM ATP and His-tagged $Cpn10_7$; closed circle, 2 mM ATP and $GroES_7$.

Mass spectrometry analysis of a mixture of various Cpn20 oligomers revealed four bands with molecular masses of 94811.9, 70980.2, 47229.6 and 23526.3 Da (Figure 4b). The signal corresponding to a tetramer was observed only at low intensity, because ionization efficiency of large molecules is low. These data provide additional strong evidence that His-tagged Cpn20 forms a tetramer. Based on the comparison of the calculated mass of the His-tagged Cpn20 (22031 Da) to that of the monomer (23527 Da), it was suggested that four DSP molecules attach to each monomeric His-tagged Cpn20.

For further confirmation of the oligomer structure of the His-tagged Cpn20, the molecular mass was also determined by native PAGE (Hedrick and Smith, 1968). Proteins were separated on 6–10% acrylamide gels at 4°C and the relative mobilities (Rm) of proteins in reference to the dye were measured. Rm varies exponentially as a function of gel concentration, and a plot of a log Rm data versus the gel concentration results in a straight line. In addition, the slope of such a plot is correlated with the molecular mass of the proteins applied (Hedrick and Smith, 1968; Figure 5). When His-tagged Cpn20 was resolved by this method, the slope was 8.6×10^{-2}, corresponding to a molecular mass of 90 kDa. The calculated value (90) divided by 22 (the molecular mass of the His-tagged Cpn20 monomer) gives 4.1, supporting that the His-tagged Cpn20 forms a tetrameric structure.

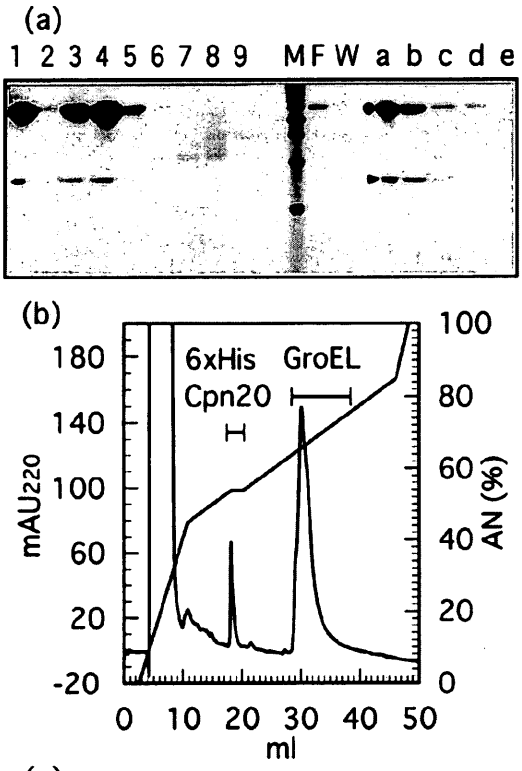
Functional analysis of a Cpn20 tetramer

Based on these experiments using the His-tagged chloroplast Cpn20 expressed in *E. coli*, we concluded that a chloroplast Cpn20 forms a tetramer. In order to study the function of such a tetramer the chaperonin-assisted protein-folding assay was performed. Denatured citrate synthase (CS) was refolded at 35°C in the presence of ATP and chaperonins. CS is not permissive for spontaneous folding at 35°C and unfolded polypeptides rapidly partition to aggregates. In this condition the recovery of CS activity depends on chaperonins (Schmidt *et al.*, 1994). GroEL was used in our assay as a partner because previous studies had suggested that Cpn20 can substitute for GroES to assist protein folding (Baneyx *et al.*, 1995). As shown in Figure 6, nearly 75% of the CS activity was recovered with GroES and His-tagged Cpn10, and 45% was recovered with His-tagged Cpn20. Doubling the amounts of Cpn20 did not change the time–course of CS folding. This indicates that the assay was run under saturating conditions for Cpn20. The sample of Cpn20 protein employed for this assay was also used for gel-filtration and native PAGE. Although there was a small peak obviously representing the monomeric form of Cpn20 (Figure 3a, top), it was not detected by CBB staining after native PAGE. Therefore, we concluded that His-tagged Cpn20 preferentially forms a tetramer under all conditions used in the experiments presented in this paper. These results also suggest that the Cpn20 tetramer can substitute for GroES and is able to function as a co-chaperonin.

The rate of CS folding with Cpn20 was slightly slower than with single-type co-chaperonins. It has been shown that spinach Cpn20 inhibits ATP hydrolysis by GroEL less effectively than GroES (Baneyx *et al.*, 1995). In our assay there were two important steps: GroEL and CS formed the binary complex and CS was released from the binary complex to exhibit its activity. This releasing step was related to inhibition of the ATPase activity of co-chap-

eronins. Therefore, it is suggested that slower folding of CS could be due to a slower release of CS from the binary complex, indicating that the formation of a complex between GroEL and Cpn20 was less efficient.

From these results the question arose whether the native chloroplast Cpn20 also forms a tetramer. The soluble fraction of chloroplasts of a tobacco wild type was analysed by gel-filtration column chromatography. Fractions of 0.5 ml were collected 14 min after the start and fractions with odd numbers were subjected to SDS–PAGE and subsequent immunoblotting with an antiserum against the His-tagged Cpn20 (Figure 3a, bottom). The protein immunoreacting with the antiserum against His-tagged Cpn20 was detected in the fractions that were eluted at 22–26 min, which is similar to the elution time for the His-tagged Cpn20 expressed in *E. coli*. These results clearly demonstrate that the native chloroplast Cpn20 has a molecular mass similar to that of the expressed Cpn20 and also forms a tetrameric structure. Moreover, Bertsch *et al.* (1992) have applied the soluble fraction of pea chloroplasts to a gel-filtration column and identified the fractions having the ability to assist GroEL in the ATP-dependent reconstitution of bacterial Rubisco (assisted protein folding assay). The native molecular mass of the protein in these fractions is about 55 kDa. As this value is similar to ours, it is suggested that the native chloroplast Cpn20 functions as a tetramer.



(c) **GroEL**

Amino acid	Amount (nmol)	Residue number	Protein (nmol)
Ala	92.95	74	1.26
Leu	52.55	42	1.25
Phe	8.90	7	1.27

6xHis Cpn20

Amino acid	Amount (nmol)	Residue number	Protein (nmol)
Ala	4.33	13	0.33
Leu	5.72	18	0.32
Phe	0.94	3	0.31

Interaction between Cpn20 and Cpn60

To determine the subunit stoichiometry of the GroEL–Cpn20 complex, we prepared the GroEL–Cpn20 complex in the presence of 1 mM ADP, as described in the Experimental procedures. SDS–PAGE of fractions after gel-filtration chromatography was performed, and a CBB-stained gel is shown in Figure 7(a). Free Cpn20 molecules were hardly detected (lanes 7, 8 and 9) in this condition. The fractions containing the complex (lanes 3 and 4) were applied to an Ni-conjugated affinity column. The Cpn20-bound GroEL only absorbed to the column and was eluted by 0.5 M imidazole (lanes a, b and c), whereas free GroEL could not absorb to the column and was then detected in the flow-through fraction (lane F). The complex was applied to a reverse-phase column, GroEL and His-tagged Cpn20 were fractionated separately (Figure 7b), and each fraction was

Figure 7. The stoichiometry of the GroEL–Cpn20 complex.
(a) The GroEL–Cpn20 complex was prepared in the presence of 1 mM ADP. The positions of GroEL and His-tagged Cpn20 are shown by a closed circle and an arrowhead, respectively (lane a). Total (lane 1) and fractions (lanes 2–9) after gel-filtration column chromatography were subjected to SDS–PAGE and subsequent staining with CBB. The fractions containing the complex (lanes 3 and 4) were applied to Ni-conjugating affinity column. The flow-through fraction (lane F), washing fraction (lane W) and fractions eluted with 0.5 M imidazole (lanes a–e) were subjected to SDS–PAGE and CBB staining. A 10-kDa protein ladder (Life Technologies, Rockville, MD) marker is shown in lane M. The lowest band corresponds to 20 kDa.
(b) Purified complex was separated by reverse-phase column chromatography. The elution profile represents the absorbance monitored at 220 nm (mAU220) and acetonitrile concentration (AN). His-tagged Cpn20 and GroEL fractions are shown by bars.
(c) The results of amino acid analysis about alanine, leucine and phenylalanine are shown. The amounts in a total fraction are indicated in the second column. The analysed volume was 2% and 20% of GroEL and Cpn20 fractions, respectively. Residue numbers were calculated from the amino acid sequence. The protein amounts were calculated from these values and are represented in the right-hand column.

subjected to amino acid analysis. The volume of GroEL and His-tagged Cpn20 used for analysis was 2% and 20% of the total fraction volume, respectively. The results of three amino acids that were recovered quantitatively well are shown in Figure 7(c). The observed amounts of amino acids were calculated from peak areas and are shown as the total amounts in each original fraction (Figure 7c, 'Amount'). These values were divided by a theoretical residue number (Figure 7c, 'Residue number') to calculate the protein amounts (Figure 7c, 'Protein'). Finally, the total amounts of GroEL and Cpn20 derived from the complex were 1.26 and 0.32 nmol, respectively, and their ratio was 14 : 3.6. These results strongly indicate that the complex contains a GroEL tetradecamer and Cpn20 tetramer.

Cpn60α and Cpn60β seem to form tetradecamers in the stroma of chloroplasts, as do GroEL. This raises the question of how the tetrameric Cpn20 is assembled with the tetradecameric Cpn60α and Cpn60β. It has been shown previously that the mobile loop region of GroES (residue 17–32) interacts with GroEL directly (Landry *et al.*, 1993). The mobile loop is a β-hairpin that extends to the interface in the GroEL–GroES complex (Xu *et al.*, 1997). Two helices of GroEL that contain mutationally identified residues responsible for peptide binding also move towards the interface and interact with a mobile loop instead of a polypeptide. Thus, it is important for GroES to shield all of the peptide binding site of GroEL. The irregular packing of the GroES heptamer suggests that its structure is flexible (Hunt *et al.*, 1996). The relative mobility of the Cpn20 tetramer was almost the same as that of the mitochondrial Cpn10 heptamer in their natural form in gel-filtration column chromatography and after artificial cross-linkage in SDS–PAGE (Figures 3 and 4a). These results strongly indicate that the rings of the Cpn10 and Cpn20 oligomer are of similar sizes. The chloroplast Cpn20 tetramer seems to be packed irregularly so that at a time seven of the eight mobile loops of the Cpn20 tetramer interact with the seven GroES binding domains of Cpn60.

Experimental procedures

Construction of a cDNA library and screening of the cDNA for the chloroplast Cpn20

The *Arabidopsis* cDNA library in *E. coli* using expression vector pBluescript™ (Stratagene, La Jolla, CA, USA) was constructed by the modified vector-primer method of Mori *et al.* (1991). Poly(A)+ RNA was prepared from 7-day-old dark-grown seedlings of *A. thaliana*, ecotype *Landsberg erecta*. Two degenerated primers, AA(A/G) CC(T/C/A/G) TC(T/C/A/G) (T/C/A/G)TT GG(T/C/A/G) AC, and A(A/G)(T/C/A/G) A(T/C)(T/C/A/G) GCC AT(T/C/A/G) AC(A/G) TC, were synthesized. Their sequences were based on that from a region of an *Arabidopsis* EST clone (accession number EMBL Z18060). The polypeptide encoded by this partial EST clone is highly homologous with the C-terminal 78 amino acids sequence of the chloroplast Cpn20 of spinach. Amplification by PCR was performed. The amplified PCR products (185 bp in length) were labelled with digoxigenin (DIG)–UTP by the random primer method and were then used as a probe for colony hybridization. Labelling and colony hybridization were performed as recommended by the supplier (Boehringer Mannheim, Tokyo, Japan).

Expression and purification of the His-tagged Cpn20 in E. coli

The cDNA fragment for Cpn20 was inserted into the expression vector pQE30 (QIAGEN, Chatsworth, CA, USA). Six histidine residues (His-tagged) located at the N-terminal of Cpn20 were derived from this vector. *Escherichia coli* cells were transformed with the construct and the expression of proteins was induced with isopropyl-B-D-thiogalacto-pyranoside (IPTG). The extracts of the sonicated *E. coli* cells were loaded onto an Ni-nitrilo-triacetic acid (Ni-NTA) column. The His-tagged Cpn20 absorbed to the resin and eluted with a 30-ml gradient of 0–0.5 M imidazole after washing.

Electrophoresis and immunoblotting

SDS–PAGE (12% acrylamide) was performed by the method of Laemmli (1970). The separated proteins on the gels were electrophoretically blotted onto cellulose nitrate membranes. Immunochemical detection was carried out by using the ECL detection system (Amersham Japan, Tokyo, Japan). An antiserum was raised in a rabbit against the His-tagged Cpn20 expressed in *E. coli*. In addition, 15% gels containing 0.15% N,N'-methylenebis acrylamide were used to separate 60–80 kDa proteins well. Native PAGE was carried out at 4°C according to Laemmli omitting SDS. After protein separation, the gel was cut at the dye-front line and stained with CBB. The relative mobilities (Rm) of proteins in reference to the dye (BPB) were measured and plotted according to the method of Hedrick and Smith (1968). The following protein markers were used to make the calibration curve: chymotrypsinogen (25 kDa), BSA (67 kDa), GroES (73 kDa), catalase (232 kDa) and ferritin (440 kDa).

Plant transformation

The β-glucuronidase gene of a binary vector pBI121Hm was replaced by a *Xba*I-*Eco*RV cDNA fragment containing the chloroplast Cpn20 cDNA. The pBI121Hm is a derivative of pBI121 and contains two drug-resistance genes, the neomycin phosphotransferase II gene and the hygromycin phosphotransferase gene. The resulting construct was introduced into *Agrobacterium tumefaciens* (strain EHA101) and used to transform *Nicotiana tobacum* var. SR1 plants by the leaf disc method (Horsch *et al.*, 1985).

Immunogold localization

Transgenic tobacco leaves were fixed, dehydrated and embedded in LR White resin (London Resin Co., Basingstoke, UK) as described previously (Nishimura *et al.*, 1993). Immunogold procedures were essentially the same as those described by Kato *et al.* (1996) except for the use of an antiserum against the His-tagged Cpn20, diluted 200-fold, and a 30-fold diluted protein A-gold (Amersham Japan, Tokyo, Japan). The sections were examined with a transmission electron microscope (1200EX; JEOL, Tokyo, Japan) at 80 kV.

Genomic Southern blotting

Total DNA was isolated from *Arabidopsis* by the cetyltrimethyl ammonium-bromide method. Two micrograms of total DNA were digested with restriction enzymes. The digests were fractionated on a 0.8% agarose gel and transferred to a nylon membrane under vacuum with 0.25 N NaOH and 1.5 M NaCl. The chloroplast Cpn20 cDNA insert was labelled with [α-^{32}P]dCTP (Amersham Japan, Tokyo, Japan) using a BcaBEST labelling kit (Takara Shuzo, Tokyo, Japan). Hybridization was performed as described by Church and Gilbert (1984). After hybridization, the membrane was sequentially washed in 2 × sodium chloride–sodium citrate buffer (SSC) and 0.1% (w/v) SDS at 42°C for 30 min, and in 0.1 × SSC and 0.1% (w/v) SDS at 65°C for 30 min. The washed membrane was exposed to an imaging plate and the radioactivity of the signals was recorded with an Analyser system (Fuji film, Tokyo, Japan).

Northern blotting

Total RNA was purified by LiCl precipitation following the extraction of RNA with ISOGEN (Nippongene, Tokyo, Japan). Ten micrograms of total RNA were electrophoresed in a 1% agarose/formaldehyde gel and transferred to a nylon membrane under vacuum with 1 N ammonium acetate. Hybridization was performed as described above for genomic Southern blotting. After hybridization the membrane was washed in 0.1 × SSC and 0.1% (w/v) SDS at 65°C for 3 h and the radioactivity of the signals was recorded as described above for genomic Southern blotting.

Gel-filtration column chromatography

The FPLC system equipped with a Superose 12 HR 10/30 column (Pharmacia Biotech, Uppsala, Sweden) was used. The column was equilibrated with 100 mM Tris–HCl, pH 7.7/10 mM $MgCl_2$/10 mM KCl (buffer A). The flow rate was 0.5 ml min^{-1} and absorbance at 280 nm was monitored. Catalase tetramer (molecular mass 232 kDa), BSA (67 kDa), RNaseA (13.7 Da) and Vitamin B_{12} (1355 Da) were used as molecular mass standards. A Hi Load 16/60 Superdex 200 pg column (Pharmacia Biotech, Uppsala, Sweden) was also used. This column was equilibrated with buffer A containing 50 mM NaCl. The flow rate was 0.8 ml min^{-1}. Thyroglobulin (669 kDa), ferritin (440 kDa), catalase, aldolase (158 kDa), BSA, ovalbumin (43 kDa), chymotrypsinogen (25 kDa) and RNaseA were used as molecular mass standards. A calibration curve defines the relationship between the elution parameter (Kav) and the log of molecular mass. The Kav values for each protein were calculated using the equation Kav = (Ve – Vo)/(Vt – Vo), where Ve = elution volume for the protein, Vo = column void volume and Vt = total bed volume.

Cross-linking with dithiobissuccinimidylpropionate (DSP)

Expressed proteins were covalently bound by a cross-linker, DSP, to analyse the oligomeric structures. DSP was dissolved in DMSO to prepare a 50-mg ml^{-1} solution. Five microlitres of DSP solution were added to 1 ml of protein solution (containing ≈ 100 μg of His-tagged Cpn20 in 0.1 M NaCl, 50 mM phosphate buffer, pH 7.2). Mixtures were incubated overnight at 4°C. Cross-linking was stopped by addition of 5 μl of 10 M ammonium acetate. The proteins were precipitated with 10% (w/v) trichloroacetic acid (TCA) and then solubilized in SDS–PAGE sample buffer. Samples were analysed by SDS–PAGE, followed by CBB staining. Because of the presence of one disulphide bond in DSP, monomers appeared after treatment with 2-mercaptoethanol (2ME).

Mass spectrometry

Positive ion MALDI-TOF-MS was performed using a Voyager Elite XL time-of-flight mass spectrometer equipped with a delayed-extraction system (PerSeptive Biosystems, Framingham, MA, USA), as described previously (Asahi *et al.*, 1997). Solutions that contained the cross-linked Cpn20 fractionated by reverse phase column chromatography were placed on the flat surface of a stainless steel plate and mixed with the matrix solution, the supernatant of a 33% acetonitrile solution saturated with sinapinic acid, and air-dried. The ions were generated by irradiating the sample area with the output of a nitrogen laser (337 nm).

Preparation of chloroplasts

Intact chloroplasts were prepared from homogenates of mature tobacco leaves by modification of the procedure described by Cline *et al.* (1985). Twenty grams of the leaves were homogenized with 150 ml of extraction medium (330 mM sorbitol/50 mM HEPES-KOH, pH 8.0/2 mM EDTA/1 mM $MgCl_2$) and filtered through three layers of cheesecloth. A crude chloroplast preparation was obtained by centrifugation at 2200 ***g*** for 30 sec. The pellet was suspended in 2 ml of extraction medium and then layered onto a discontinuous gradient of Percoll. The gradient was composed of 90% (8 ml), 40% (12 ml), 30% (6 ml) and 10% (6 ml) Percoll (v/v) in extraction medium. The fraction that contained intact chloroplasts was located between 90% and 40% Percoll after centrifugation for 20 min at 8800 ***g***. This fraction was diluted fourfold with extraction medium and centrifuged at 2200 ***g*** for 30 sec. The chloroplast pellet was resuspended in 0.01% Triton X-100 and sonicated to obtain the soluble extracts. The soluble extracts of chloroplasts was then applied to a gel-filtration column.

Chaperonin-assisted protein folding assay

A chaperonin-assisted protein assay was performed as described by Schmidt *et al.* (1994). Citrate synthase (CS, 15 μM) was denatured in a solution containing 6 M GuHCl/100 mM Tris–HCl, pH 8.0/20 mM dithiothreitol for at least 1 h at room temperature. Denatured CS was rapidly diluted to a concentration of 150 nM into a solution containing 50 mM Tris–HCl, pH 7.7/10 mM $MgCl_2$/10 mM KCl/225 nM GroEL tetradecamer at 0°C. The temperature was adjusted to 35°C and 2 mM ATP and 300 nM co-chaperonin oligomer were added. After various times of incubation aliquots were removed and assayed for CS activity, essentially as described by Srere (1969). GroEL and GroES were purchased from Takara Shuzo (Tokyo, Japan). The amount of co-chaperonin was determined by a protein assay (BioRad Laboratories, Hercules, CA, USA).

Preparation of GroEL–Cpn20 complexes

The GroEL–Cpn20 complex was formed by incubating $GroEL_{14}$ (2.1 μM) and His-tagged $Cpn20_4$ (1.3 μM) in 100 mM Tris–Cl, pH 7.7/10 mM $MgCl_2$/10 mM KCl/50 mM NaCl/1 mM ADP (buffer B) at room temperature for 1.5 h. First, the complex and free GroEL were separated from free His-tagged Cpn20 by gel-filtration column chromatography in the presence of 1 mM ADP. Second, this fraction was applied onto an Ni-conjugated affinity column (HiTrap Chelat-

ing column; Pharmacia Biotech, Uppsala, Sweden) in the presence of 1 mM ADP. Free GroEL was washed out and the complex was eluted with 500 mM imidazole. GroEL and His-tagged Cpn20 were fractionated separately by reverse phase column chromatography (Sephasil Protein C4 5 μm ST 4.6/250 column; Pharmacia Biotech, Uppsala, Sweden) for amino acid analysis.

Amino acid analysis

Gas-phase hydrolysis with 200 μl of 6 N HCl containing 0.2% phenol was carried out at 110°C for 24 h. The hydrolysates were analysed in an Hitachi L-8500 A amino acid analyser.

Acknowledgements

The authors thank Dr H. Mori (Nagoya University) for the construction of the cDNA library, and Ms S. Ohsawa (The Center for Analytical Instruments, NIBB) for amino acid analysis. They also thank Dr R.R. Theimer (Bergische University, Germany) for critical reading of the manuscript. This work was supported in part by a grant from the Research for the Future Program of the Japanese Society for the Promotion of Sciences (JSPS-RFTF96 I00407), and by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan (09274101, 09274103, 09440271).

References

Asahi, M., Fujii, J., Takao, T., Kuzuya, T., Hori, M., Shimonishi, Y. and Taniguchi, N. (1997) The oxidation of selenocysteine is involved in the inactivation of glutathione peroxidase by nitric oxide donor. *J. Biol. Chem.* **272**, 19152–19157.

Baneyx, F., Bertsch, U., Kalbach, C.E., Vies, S.M.v.d., Soll, J. and Gatenby, A.A. (1995) Spinach chloroplast cpn21 co-chaperonin possesses two functional domains fused together in a toroidal structure and exhibits nucleotide-dependent binding to plastid chaperonin 60. *J. Biol. Chem.* **270**, 10695–10702.

Bertsch, U. and Soll, J. (1995) Functional analysis of isolated cpn10 domains and conserved amino acid residues in spinach chloroplast co-chaperonin by site-directed mutagenesis. *Plant Mol. Biol.* **29**, 1039–1055.

Bertsch, U., Soll, J., Seetharam, R. and Viitanen, P.V. (1992) Identification, characterization, and DNA sequence of a functional 'double' groES-like chaperonin from chloroplasts of higher plants. *Proc. Natl Acad. Sci. USA*, **89**, 8696–8700.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L. and Sigler, P.B. (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature*, **371**, 578–586.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci. USA*, **81**, 1991–1995.

Cline, K., Werner-Washburne, M., Lubben, T.H. and Keegstra, K. (1985) Precursors to two nuclear-encoded chloroplast proteins bind to the outer envelope membrane before being imported into chloroplasts. *J. Biol. Chem.* **260**, 3691–3696.

Cloney, L.P., Wu, H.B. and Hemmingsen, S.M. (1992) Expression of plant chaperonin-60 genes in *Escherichia coli*. *J. Biol. Chem.* **267**, 23327–23332.

Cole, K.P., Blakeley, S.D. and Dennis, D.T. (1994) Isolation of a full-length cDNA encoding *Brassica napus* mitochondrial chaperonin-60. *Plant Physiol.* **105**, 451.

Fossati, G., Lucietto, P., Giuliani, P., Coates, A.R., Harding, S., Cölfen, H., Legname, G., Chan, E., Zaliani, A. and Mascagni, P. (1995) Mycobacterium tuberculosis chaperonin 10 forms stable tetrameric and heptameric structures. *J. Biol. Chem.* **270**, 26159–26167.

Hartl, F.U. (1996) Molecular chaperones in cellular protein folding. *Nature*, **381**, 571–580.

Hedrick, J.L. and Smith, A.J. (1968) Size and charge isomer separation and estimation of molecular weights of proteins by disc gel electrophoresis. *Arch. Biochem. Biophys.* **123**, 155–164.

Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W. and Ellis, R.J. (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature*, **333**, 330–334.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) A simple and general method for transferring genes into plants. *Science*, **227**, 1229–1231.

Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L. and Deisenhofer, J. (1996) The crystal structure of the GroES co-chaperonin at 2.8Å resolution. *Nature*, **379**, 37–45.

Kato, A., Hayashi, M., Kondo, M. and Nishimura, M. (1996) Targeting and processing of a chimeric protein with the amino-terminal presequence of the precursor to glyoxysomal citrate synthase. *Plant Cell*, **8**, 1601–1611.

Koumoto, Y., Tsugeki, R., Shimada, T., Mori, H., Kondo, M., Hara-Nishimura, I. and Nishimura, M. (1996) Isolation and characterization of a cDNA encoding mitochondrial chaperonin 10 from *Arabidopsis thaliana* by functional complementation of an *Escherichia coli groES* mutant. *Plant J.* **10**, 1119–1125.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.

Landry, S.J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C. and Gierasch, L.M. (1993) Characterization of a functionally important mobile domain of GroES. *Nature*, **364**, 255–258.

Madueño, F., Napier, J.A. and Gray, J.C. (1993) Newly imported Rieske iron-sulfur protein associates with both Cpn60 and Hsp70 in the chloroplast stroma. *Plant Cell*, **5**, 1865–1876.

Martel, R., Cloney, L.P., Pelcher, L.E. and Hemmingsen, S.M. (1990) Unique composition of plastid chaperonin-60: α and β polypeptide-encoding genes are highly divergent. *Gene*, **94**, 181–187.

Mori, H., Takeda-Yoshikawa, Y., Hara-Nishimura, I. and Nishimura, M. (1991) Pumpkin malate synthase: cloning and sequencing of the cDNA and Northern blot analysis. *Eur. J. Biochem.* **197**, 331–336.

Nishimura, M., Takeuchi, Y., De Bellis, L. and Hara-Nishimura, I. (1993) Leaf peroxisomes are directly transformed to glyoxysomes during senescence of pumpkin cotyledons. *Protoplasma*, **175**, 131–137.

Rospert, S., Junne, T., Glick, B.S. and Schatz, G. (1993) Cloning and disruption of the gene encoding yeast mitochondrial chaperonin 10, the homolog of *E. coli* groES. *FEBS Lett.* **335**, 358–360.

Ryan, M.T., Hoogenraad, N.J. and Høj, P.B. (1994) Isolation of a clone specifying rat chaperonin 10, a stress-inducible mitochondrial matrix protein synthesised without a cleavable presequence. *FEBS Lett.* **337**, 152–156.

Schlicher, T. and Soll, J. (1996) Molecular chaperones are present in the thylakoid lumen of pea chloroplasts. *FEBS Lett.* **379**, 302–304.

Schmidt, M., Buchner, J., Todd, M.J., Lorimer, G.H. and Viitanen, P.V. (1994) On the groES in the chaperonin-assisted folding reaction. *J. Biol. Chem.* **269**, 10304–10311.

Srere, P.A. (1969) Citrate synthase. *Methods Enzymol.* **13**, 3–11.
Todd, M.J., Viitanen, P.V. and Lorimer, G.H. (1994) Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. *Science*, **265**, 659–666.
Tsugeki, R. and Nishimura, M. (1993) Interaction of homologues of Hsp70 and Cpn60 with ferredoxin-$NADP^+$ reductase upon its import into chloroplasts. *FEBS Lett.* **320**, 198–202.
Weissman, J.S., Hohl, C.M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H.R., Fenton, W.A. and Horwich, A.L. (1995) Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. *Cell*, **83**, 577–587.
Xu, Z., Horwich, A.L. and Sigler, P.B. (1997) The crystal structure of the asymmetric GroEL-GroES-$(ADP)_7$ chaperonin complex. *Nature*, **388**, 741–750.
Zabaleta, E., Oropeza, A., Jimenez, B., Salerno, G., Crespi, M. and Herrera-Estrella, L. (1992) Isolation and characterization of genes encoding chaperonin 60β from *Arabidopsis thaliana*. *Gene*, **111**, 175–181.

The nucleotide sequence of *Arabidopsis* Cpn20 cDNA has been submitted to the EMBL/GenBank/DDBJ nucleotide sequence databases under accession number AB007130.

Chloroplasts Have a Novel Cpn10 in Addition to Cpn20 As Co-chaperonins in *Arabidopsis thaliana*

Yasuko Koumoto[1], Tomoo Shimada[1], Maki Kondo[2], Ikuko Hara-Nishimura[1] and Mikio Nishimura[2]*

1 Department of Botany, Graduate School of Science, Kyoto University, Kyoto, 606-8502 Japan

2 Department of Cell Biology, National Institute for Basic Biology, Okazaki, 444-8585 Japan

*For Correspondence; Mikio Nishimura

National Institute for Basic Biology

Okazaki, 444-8585 Japan

Phone, +81 564-55-7500

Fax, +81 564-55-7505

e-mail, mikosome@nibb.ac.jp

Running title: Chloroplast Cpn10 in Arabidopsis

Summary

Previously, we characterized a mitochondrial (Cpn10) and a chloroplast co-chaperonins (Cpn20) from *Arabidopsis thaliana* [Koumoto *et al.* (1996) *Plant J.* 10, 1119-1125; Koumoto *et al.* (1999) *Plant J.* 17, 467-477]. Here, we report a third co-chaperonin. The cDNA was 603 bp long encoding a protein of 139 amino acids. From a sequence analysis, the protein was predicted to have one Cpn10 domain with an amino-terminal extension that might work as a chloroplast transit peptide. This novel Cpn10 was confirmed to be localized in chloroplasts, and we refer to it as chloroplast Cpn10 (chl-Cpn10). The phylogenic tree that was generated with amino acid sequences of other co-chaperonins, indicates that chl-Cpn10 is highly divergent from the others. In the GroEL-assisted protein-folding assay, about 30% of the substrate were refolded with chl-Cpn10, indicating that chl-Cpn10 works as a co-chaperonin. A Northern blot analysis revealed that mRNA for chl-Cpn10 is accumulated in leaves and stems, but not in roots. In germinating cotyledons, the accumulation of chl-Cpn10 was similar to that of chloroplastic proteins and accelerated by light. It was proposed that two kinds of co-chaperonins, Cpn20 and chl-Cpn10 work in the chloroplast independently.

Introduction

Chaperonins are proteins that play a vital role in protein folding in eukaryotic and prokaryotic cells. They are generally divided into two groups (1). Group I chaperonins are localized in the stroma of chloroplasts, the matrix of mitochondria and eubacteria, and group II chaperonins are found in the eukaryotic cytosol and archaebacteria. One of the differences between the two groups is the necessity of co-chaperonins for their function. Group I chaperonins work together with co-chaperonins, while group II chaperonins work alone.

The Chaperonin (GroEL) and co-chaperonin (GroES) of *Escherichia coli* have been characterized in detail. They are heat-shock proteins that are also required for viability under normal condition (2). GroEL is composed of two heptameric rings, each consisting of seven 57 kDa subunits. Two stacked rings of the GroEL heptamer form a central cavity that captures incompletely folded proteins. The co-chaperonin GroES, a dome-shaped ring consisting of seven 10 kDa subunits, binds to GroEL. GroES has been shown to increase the co-operativity of the ATPase activity of GroEL (3). The crystal structures of the GroEL tetradecamer (4), the GroES heptamer (5), the GroEL-GroES complex (6) and GroEL-peptide complex (7) have been characterized.

While the structures and basic mechanism of reaction have been well defined, little is known about the roles of the chaperonins *in vivo*. From the analysis of a temperature-sensitive GroEL mutant strain of *E. coli*, it has been shown that about 30% of newly translated polypeptides fold via GroEL (8). Moreover, it has been shown that about 10-15% of cytoplasmic proteins of *E. coli* cells interact with GroEL (9). These results suggested that the majority of proteins fold without the help of chaperonins *in vivo*. However, it is a fact that the folding of some proteins severely depends on chaperonins (10). From import experiments of temperature-sensitive mutants of yeast mitochondrial chaperonin (Cpn60) and co-chaperonin (Cpn10), it was shown that some proteins form aggregates in the absence of either Cpn60 or Cpn10. Especially, the folding of a newly imported Cpn60 itself severely depends on Cpn10.

The chaperonin system in the chloroplast stroma of higher plants is unique.

The stromal chaperonin oligomer consists of two isoforms, Cpn60α and Cpn60β (11) and stromal co-chaperonin (Cpn20) is comprised of two GroES-like tandem domains (12). Schlicher and Soll (13) showed that the thylakoid lumen contains a Cpn10 homologue that is recognized by a peptide specific antiserum raised against a peptide of the stromal Cpn20. The reported size of the lumenal homologue on SDS-PAGE is about 10-12 kDa. On the other hand, in a proteomics study on the chloroplast of pea performed recently, Peltier *et al.* (14) identified Hsp70, Cpn60α and Cpn20, on the two-dimensional electrophoresis map with thylakoid lumenal proteins. The explanation for the presence of stromal chaperones in the lumenal fraction was that they are bound to the stromal side of the thylakoid membrane. They supposed that the 10-12 kDa protein recognized by the antiserum against Cpn20 was a degradative product of Cpn20 and not a Cpn10 homologue. Thus, we performed studies to examine whether a Cpn10 homologue exists in the chloroplast or not. In this paper, we report the existence of a novel Cpn10 homologue in the chloroplast. Sequence analysis showed that chloroplast Cpn10 was divergent from the other co-chaperonins, suggesting that it might have evolved to have a special function.

Experimental procedures

Expression and purification of the His-tagged Cpn10 homologue in E. coli

The cDNA fragment for the deduced mature region of chl-Cpn10, residues 50-139, was inserted into the expression vector pQE30 (QIAGEN, Chatsworth, CA, U.S.A.). Six histidine residues (His-tagged) located at the N-terminal of chl-Cpn10 were derived from this vector. *E. coli* was transformed with the construct and the expression of proteins was induced with isopropyl-β-D-thiogalactopyranoside (IPTG).

Chaperonin-assisted protein folding assay

The chaperonin-assisted protein folding assay was performed as described by Schmidt *et al.* (15). Citrate synthase (CS) was denatured at a concentration of 15 μM in a solution containing 6 M GuHCl / 100 mM Tris-Cl, pH8.0 / 20 mM dithiothreitol for at

least 1 hour at room temperature. Denatured CS was rapidly diluted to a concentration of 150 nM into a solution containing 50 mM Tris-Cl, pH7.7 / 10 mM $MgCl_2$ / 10 mM KCl / 225 nM GroEL tetradecamer at 0ÚC. The temperature was adjusted to 35ÚC and 2 mM ATP and 300 nM co-chaperonin oligomer was added. After various lengths of incubation, aliquots were removed and assayed for CS activity. The assay for CS was essentially as described (16). GroEL and GroES were purchased from Takara (Tokyo, Japan). The content of co-chaperonin was determined by a protein assay (BioRad Laboratories, Hercules, CA, U.S.A.).

Immunogold localization

Transgenic Arabidopsis cotyledons were fixed, dehydrated and embedded in LR White resin (London Resin Co., Basingstoke, UK) as described previously (17). Immunogold procedures were essentially the same as those described by Kinoshita *et al.* (18), except for the use of antiserum against the His-tagged chl-Cpn10 diluted 50-fold. 15 nm protein A-gold (Amersham Japan, Tokyo, Japan) was diluted 100-fold to use. An antiserum was raised in a rabbit against the His-tagged chl-Cpn10 expressed in *E. coli*. The sections were examined with a transmission electron microscope (1200EX; JEOL, Tokyo, Japan) at 80 kV.

Preparation of chloroplasts

Intact chloroplasts were prepared from homogenates of mature tobacco leaves as described in previous paper (19). The chloroplast pellets equivalent to 130 μg chlorophyll were lysed in 150 μl of 10 mM Hepes-KOH, pH6.0 / 10 mM $MgCl_2$. After 5 minutes on ice, the membrane and the soluble fractions were separated by centrifugation at 4000 x **g** for 10 minutes.

Northern blotting

Total RNA was purified by LiCl precipitation following the extraction of RNA with ISOGEN (Nippongene, Tokyo, Japan). Twenty microgram of total RNA was electrophoresed on a 1% agarose/formaldehyde gel and transferred to a nylon membrane

under vacuum with 1 N ammonium acetate. Hybridization was performed as described by Church and Gilbert (20). After hybridization, the membrane was sequentially washed in 2 × sodium chloride-sodium citrate buffer (SSC) and 0.1% (w/v) SDS at room temperature for 30 minutes twice and in 0.1 × SSC and 0.1% (w/v) SDS at 60ÚC for 30 minutes twice. The washed membrane was exposed to an imaging plate and radioactivity of signals were detected with an analyzer system (Fuji Film, Tokyo, Japan).

Electrophoresis and immunoblotting

SDS-PAGE (15% acrylamide) was performed by the method of Laemmli (21). The separated proteins on the gels were electrophoretically blotted onto polyvinylidene difluoride membranes (Nihon Millipore Ltd., Tokyo, Japan). Immunochemical detection was performed by the ECL detection system (Amersham Japan, Tokyo, Japan). After detection, the blotted membrane was rinsed once in water and stained in 0.25% Coomassie Brilliant Blue (CBB) solution for 1minutes The stained membrane was washed once in 50% methanol until the protein bands appeared and were dried.

Plant transformation

To generate overproducing transgenic plants, The β-glucuronidase gene of a binary vector pBI121Hm was replaced by a SalI-XbaI cDNA fragment derived from an Arabidopsis EST clone (GenBank accession no. T44192). The pBI121Hm is a derivative of pBI121 and contains two drug resistance genes, the neomycin phosphotransferase II gene and the hygromycin phosphotransferase gene. The resulting construct was introduced into *Agrobacterium tumefaciens* (strain EHA101) and used to transform *Arabidopsis thaliana* ecotype Columbia plants by the *in planta* method (22).

Quantitative RT-PCR

First strand cDNA was synthesized using a Superscript^{TM} preamplification system for first strand cDNA synthesis kit (Gibco BRL, Rockville, MD) from total RNA. The oligo (dT) primer was used for reverse transcription. Quantitative PCR assays were performed by an ABI Prism 7700 sequence detection system (PE Biosystems Japan,

Tokyo, Japan). This system is capable of detecting PCR products as they accumulate during PCR. For detection, it is necessary to prepare the fluorescent DNA probe that is specific to the PCR target sequence flanked by PCR primers. To select the PCR primers and the fluorescent probes, we used the program, Prime Express^{TM} (PE Biosystems Japan, Tokyo, Japan). All PCR procedures were carried out with 40 cycles of denaturation at 95ÚC for 15 seconds with annealing and extension at 60 ÚC for 1 minute according to the supplier.

Results & Discussion

Sequence analysis

In *Arabidopsis thaliana*, cDNAs for mitochondrial Cpn10 and chloroplast Cpn20 have already been cloned (19, 23, 24). Using amino acid sequences of the full-length of the former and the mature region of the latter, we searched the EST database for another homologous cDNA. As a result, some EST clones were found to have the amino-terminal conserved motif of Cpn10, PXX(D/N)(K/R) (Figure 1a), but showed low similarity to the above two clones as a whole. The sequence of the longest EST clone (GenBank accession no. T44192) was determined. The cDNA was 603 bp long encoding a protein of 139 amino acids (Figure 1a). The protein was predicted to be a homologue of Cpn10 containing an amino terminal extension from sequence analysis. This extension was predicted to be a chloroplast transit peptide using PSORT program (http://www.nibb.ac.jp/). Therefore, the Cpn10 homologue is referred to as chloroplast Cpn10 (chl-Cpn10). Moreover, we found tomato and soybean EST clones that exhibit high sequence homology with chl-Cpn10 of Arabidopsis (Figure 1b). This indicates that chl-Cpn10 exists in various plants. Genome sequencing revealed another clone of chl-Cpn10 in Arabidopsis (25). The amino acid sequences of the two Arabidopsis clones showed 72% identity.

The phylogenic tree was generated with amino acid sequences of other co-chaperonins including bacterial GroES (Figure 2). The amino-terminal conserved motif of GroES starts from the 5th proline. Because it is expected that sequences with similar length were aligned more accurately, the amino acid sequence from Thr50 to Glu139 of

chl-Cpn10, in which the conserved motif also starts from the 5th proline same as GroES, was used for alignment. As shown in Figure 2, at chl-Cpn10 is highly divergent from other co-chaperonins. For more detailed analysis, chl-Cpn10 was compared with GroES (Figure 1c). The crystal structure of GroES has 9 β-sheets and the mobile loop (5). The mobile loop is the region that interacts with GroEL. The fourth and fifth β-sheets compose the roof structure of the GroES heptameric dome. When both the 8residue gap between Lys106 and Arg107 of chl-Cpn10 and 2-residue gap in the mobile loop region of GroES were inserted, identical residues were found in the corresponding regions of third, sixth and ninth β-sheet of GroES (Figure 1c). Thus, chl-Cpn10 might lack one domain and the oligomer of chl-Cpn10 might have a smaller or no roof structure.

Gp31 is a co-chaperonin encoded by bacteriophage T4 and mediates the folding and assembly of the T4 major capsid protein, Gp23, with GroEL. Because T4 cannot propagate in the absence of Gp31, GroES cannot substitute for Gp31 (21). From its crystal structure, the roof structure of Gp31 was found to be completely absent, leaving a cavity at least 16 Å in diameter (26). Due to the absence of the roof structure of Gp31, the complexes between GroEL and Gp31 might have a bigger substrate-binding cavity than GroEL and GroES. The molecular mass of Gp23, 55 kDa, is almost the limit of size to interact with GroEL. It has been speculated that Gp31 evolved to expand the substrate-binding cavity of GroEL and in order to mediate the folding of Gp23 more easily. It is an exciting speculation that chl-Cpn10 has its own specific substrates and evolved with a specialized function. In fact, chl-Cpn10 has a different activity from Cpn20 as a co-chaperonin of GroEL on the refolding of denatured protein as follows.

Activity of chl-Cpn10 as a co-chaperonin

The chaperonin-assisted protein-folding assay was performed to study whether chl-Cpn10 works as a co-chaperonin. The polypeptide containing residues 50-139 of chl-Cpn10 was expressed in *E. coli*. The segment containing the 6-histidine residues (His-tag) was fused to the N-terminus. Since the expressed His-tagged chl-

Cpn10 was insoluble, it was immobilized on a Ni-coupling column after being solubilized with 8 M urea and renaturated by a gradual decrease of urea from 6 M to 1 M, then eluted by imidazole. The purified His-tagged chl-Cpn10, judged from the CBB staining pattern (Figure 3a), was used for assays. As a substrate, citrate synthase (CS) was used. At 35ÚC, denatured CS is not permissive for spontaneous folding and the recovery of CS activity depends on chaperonins (15). GroEL was used as a chaperonin. As shown in Figure 3b, about 30% of the CS activity were recovered with the His-tagged chl-Cpn10. By doubling the amounts of chl-Cpn10, the recovery of CS activity was increased slightly. These results confirmed that the His-tagged Cpn10 homologue could interact with GroEL and work as a co-chaperonin

In this assay, the authentic co-chaperonin, GroES, recovered about 70% activity. The previous study showed that the activity of mitochondrial Cpn10 containing His-tag at its N-terminus was the same as that of GroES (19). This indicates that the N-terminal His-tag has no effect on His-tagged mitochondrial Cpn10 in this assay, and His-tag seemed not to prevent co-chaperonins from binding to GroEL. It was shown that about 55% of the CS activity were recovered with His-tagged Cpn20 (Figure 3). The activity of Cpn20 was higher than that of chl-Cpn10, although chl-Cpn10 is the same single-type co-chaperonin as GroES. It was suggested that the low activity of chl-Cpn10 was caused by the sequence specificity. As the length of the mobile loop of chl-Cpn10 is expected to be longer than that of GroES, it seems that chl-Cpn10 is not able to bind GroEL efficiently.

Localization of chl-Cpn10

To confirm the localization of chl-Cpn10, immunogold labeling was performed. We have previously revealed by immunogold labeling that mitochondrial Cpn10, which was expressed in a transgenic plant, is exclusively localized in mitochondria (23). As endogenous chl-Cpn10 was hardly detectable by immunogold labeling, the transgenic Arabidopsis harboring chl-Cpn10 cDNA was prepared. The chl-Cpn10 was expressed under regulation of the 35S promoter from cauliflower mosaic virus. An immunoblot analysis of wild-type and transgenic Arabidopsis using

rosette leaves was shown in Figure 4a. Because the immunoreactive bands in both lanes showed almost the same size, it is expected that the overexpressed chl-Cpn10 was imported into chloroplasts and processed correctly.

Seedlings of transgenic Arabidopsis, which were grown in the dark for 4 days, were transferred and kept in continuous light for 1 day. Cotyledons became green within 1 day of illumination and their chloroplasts had stacked thylakoid membranes. The gold particles for chl-Cpn10 were exclusively restricted to a chloroplast (Figure 4b) and most of them were in contact with thylakoid membranes (Figure 4c). Fractionation analysis was performed to examine whether chl-Cpn10 was associated with thylakoid membrane. Because it was difficult to isolate chloroplasts from Arabidopsis, intact chloroplasts were isolated from mature tobacco leaves. After disruption by osmotic shock, the insoluble precipitates containing thylakoid membrane and the soluble fraction were separated. Immunoblot analysis of these two fractions was shown in Figure 5. The immunoreactive band for chl-Cpn10 was detected in the soluble fraction. Cpn20, which was used as a stromal control, was concentrated in the soluble fraction. These results indicated that chl-Cpn10 was not associated with thylakoid membrane.

Analysis of expression patterns

The expression pattern of chl-Cpn10 gene was also examined. Figure 6a shows the Northern blotting of RNA extracted from roots, stems and leaves of mature Arabidopsis plants. Because of the weak signal of chl-Cpn10, 20 μg of total RNA was used. The amount of ribosomal RNA in each lane was almost the same. The mRNA for chl-Cpn10 was present in leaves and stems, but the signal was not detected in roots. Cpn20 was observed in these three organs (19). These results suggested that chl-Cpn10 and Cpn20 work independently.

Next, the effect of heat treatment on the accumulation of chl-Cpn10 mRNA was examined. For detailed study, we performed quantitative RT-PCR with a fluorescent probe, which is specific to the target PCR product. First, the total RNA was prepared from Arabidopsis plants incubated at 35ÚC for various lengths of time. Five μg of RNA was used for reverse transcription with an oligo (dT) primer. Because almost

the same results were observed with duplicates, it seems to be unnecessary that the efficiency of reverse transcription was taken into consideration. One gene, APETARA 2 (AP2, 27) was used as a control, because the amount of AP2 mRNA was constant during heat treatment. Figure 6b shows the amount of chl-Cpn10 mRNA that was normalized by the amount of AP2 mRNA. Heat treatment did not change the amount of chl-Cpn10 mRNA drastically. This result was distinct from those of both mitochondrial Cpn10 and chloroplast Cpn20 whose mRNAs showed nearly tenfold and fivefold increases by heat treatment, respectively. This also indicated that chl-Cpn10 belongs to a divergent class of co-chaperonin.

Chaperones in chloroplasts

To investigate the differences between chl-Cpn10 and Cpn20, the accumulation pattern of chl-Cpn10 protein during germination and greening was examined. Arabidopsis seeds were germinated in the dark. Cotyledons were yellow during the incubation in the dark and became fully green within 1 day of exposure to light. In cells exposed to light, proteins involved in photosynthetic electron transfer, such as LHCII, were rapidly accumulated (Figure 7, bottom). One of the major stromal proteins involved in photosynthetic carbon metabolism, RuBisCO, was accumulated gradually in the dark, especially after a 5-day incubation. A comparison of the amounts of RuBisCO in the lane '5' and that in '4D1L', indicated that the accumulation was also accelerated by light. Accumulation of chl-Cpn10 was observed after a 5-day incubation in the dark and accelerated by light similar to RuBisCO (Figure 7, top). The amount of chl-Cpn10 in a 7-day dark-grown (7D) cotyledon was comparable to that in a light-exposed cotyledon. Transgenic Arabidopsis that overexpresses chl-Cpn10 was also germinated and grown for 7 days in the dark. Immunogold labeling of transgenic Arabidopsis showed that chl-Cpn10 was localized in etioplasts of a 7D-cotyledon (Figure 4d).

In contrast to chl-Cpn10, accumulation of Cpn20 was observed at an early stage of germination. This also suggested that chl-Cpn10 and Cpn20 work independently. Cpn60α and Cpn60β had been shown to bind to immobilized Cpn20

(28). As these Cpn60s were originally referred to as RuBisCO binding proteins, it is known that one of their substrates is RuBisCO. Therefore, it is thought that Cpn20 facilitates the folding of RuBisCO with Cpn60α and Cpn60β. It might be reasonable that Cpn20 protein seemed to accumulate slightly faster than RuBisCO. By analogy, chl-Cpn10 might mediate the folding of the protein that accumulates later.

From a search of Arabidopsis genomic sequence, we found two Cpn60α genes and four Cpn60β genes. One of the two Cpn60α and three of four Cpn60β may actually be expressed, judging from the presence of EST clones. As chl-Cpn10 and Cpn20 are thought to work independently, there is a possibility that they regulate a specific partner, Cpn60. It had been reported that Cpn60α and Cpn60β were expressed in *E. coli* and were characterized (29). This report indicated that Cpn60β could form oligomer by itself, although Cpn60α could not. Moreover, it also indicated that a homo-oligomer of Cpn60β works with only mitochondrial Cpn10, in contrast to a hetero-oligomer that is compatible with various co-chaperonins. These results suggested the possibility that one Cpn60β homologue could form its homo-oligomer and function with its specific co-chaperonin. It is possible that a chaperonin system containing chl-Cpn10 and a Cpn60β homologue exists in chloroplasts in addition to one containing Cpn20, Cpn60α and Cpn60β.

Acknowledgment

The authors thank Dr. Hiroshi Hayashi (NIBB) for providing the sample of Arabidopsis. This work was supported in part by a grant from the Research for the Future Program of the Japanese Society for the Promotion of Sciences (JSPS-RFTF96L00407).

References

1. Braig, K. (1998) *Current Opinion in Structural Biology*. **8**, 159-165.

2. Fayet, O., Ziegelhoffer, T. and Georgopoulos, C. (1989) *J. Bacteriol.* **171,** 1379-1385.

3. Todd, M.J., Viitanen, P.V. and Lorimer, G.H. (1994) *Science* **265**, 659-666.

4. Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L. and Sigler, P.B. (1994) *Nature* **371**, 578-586.

5. Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L. and Deisenhofer, J. (1996) *Nature* **379**, 37-45.

6. Xu, Z., Horwich, A.L. and Sigler, P.B. (1997) *Nature* **388**, 741-750.

7. Chen L. and Sigler, P.B. (1999) *Cell* **99**, 757-768.

8. Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, I. N. and Furtak, K. (1993) *Cell* **74**, 909-917.

9. Ewalt, K. L., Hendrick, J., P., Houry, W., A. and Hartl F., U. (1997) *Cell* **90**, 491-500.

10. Dubaquié, Y., Looser, R., Fünfschilling, U., Jenö, P. and Rospert, S. (1998) *EMBO J.* **20**, 5868-5876.

11. Nishio K, Hirohashi T, Nakai M. (2000) *Biochem Biophys Res Commun.* **266,** 584-587.

12. Bertsch, U., Soll, J., Seetharam, R.and Viitanen, P.V. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**,8696-8700.

13. Schlicher, T. and Soll, J. (1996) *FEBS Lett.* **379**, 302-304.

14. Peltier, J-B., Friso, G., Kalume, D.E., Roepstorff, P., Nilsson, F., Adamska, I. and Wijk. K.J.v. (2000) *Plant J.* **12,** 319-341.

15. Schmidt, M., Buchner, J., Todd, M.J., Lorimer, G.H. and Viitanen, P.V. (1994) *J. Biol. Chem.* **269**, 10304-10311.

16. Srere, P.A. (1969) *Methods Enzymol.* **13**, 3-11.

17. Nishimura, M., Takeuchi, Y., De Bellis, L. and Hara-Nishimura, I. (1993) *Protoplasma* **175**, 131-137.

18. Kinoshita, T., Yamada, K., Hiraiwa, N., Kondo, M., Nishimura, M. and Hara-Nishimura, I. (1999) *Plant J.* **19**,43-53.

19. Koumoto, Y., Shimada, T., Kondo, M., Takao, T., Shimonishi, Y., Hara-Nishimura, I. and Nishimura, M. (1999) *Plant J.* **17**, 467-477.

20. Church, G.M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991-1995.

21. Laemmli, U.K. (1970) *Nature* **227**, 680-685.

22. Bethtold, N., Ellis, J. & Pelletier, G. (1993) *C. R. Acad. Sci. Ser. III Sci. Vie.* **316,** 1194-1199.

23. Koumoto, Y., Tsugeki, R., Shimada, T., Mori, H., Kondo, M., Hara-Nishimura, I. and Nishimura, M. (1996) *Plant J.* **10**, 1119-1125.

24. Hirohashi T, Nishio K, Nakai M. (1999) *Biochim Biophys Acta.* **1429**,512-515.

25. The Arabidopsis Genome Initiative (2000) *Nature* 408, 796-815.

26. Hunt, J. F., van der Vies, S.,M., Henry, L. and Deisenhofer, J. (1997) *Cell* **90**, 361-371.

27. Jofuku, K.D., den Boer, B.G., Van Montagu, M. and Okamuro, J.K. (1994) *Plant Cell* **6**, 1211-1225.

28. Baneyx, F., Bertsch, U., Kalbach, C.E., Vies, S.M.v.d., Soll, J. and Gatenby, A.A. (1995) *J. Biol. Chem.* **270**, 10695-10702.

29. Dickson, R., Weiss, C., Howard, R.J., Alldrick, S.P., Ellis, R.J., Lorimer, G., Azem, A. and Viitanen, P.V. (2000) *J. Biol. Chem.* **275**, 11829-11835.

Figure legends

Figure1. Nucleotide and deduced amino acid sequences of a cDNA for the chl-Cpn10 in Arabidopsis.

(a) Nucleotides are numbered from the 5' to the 3' region. The deduced amino acid sequence is numbered from the first methionine residue. The asterisk indicates a termination codon. The amino-terminal PXX(D/N)(K/R) motif is boxed.

(b) Chl-Cpn10 identified in this paper was aligned with other plant homologues. The accession numbers are: Arabidopsis2, AL139659 (88880-90146); tomato, AW455289; soybean, AW733561.

(c) The amino acid sequence of deduced mature region of chl-Cpn10 was compared with that of *E. coli* GroES. Structural features of GroES, which have been observed

crystallographically, are cited from the reference (5) and shown below sequences. Arrows indicate β-strands. Identical residues are boxed.

Figure 2. Relationship among amino acid sequences of co-chaperonins

The UPGMA program (GeneWorks; Intelligenetics, Moutain View, CA) was used to calculate the relative lengths of branches. The numbers indicate the validity of the branch point. A mitochondrial homologue was referred to as mCpn10. The accession numbers are: Pseudomonas-GroES, P30720; Escherichia-GroES, U00096; Staphylococcus-GroES, Q08841; Bacillus-GroES, P28599; Rickettsia-GroES, P80469; Agrobacterium-GroES, X68263; Synechococcus-GroES, M58751; Spinach-Cpn20, M87646; Brassica-mCpn10, U65890; Arabidopsis-mCpn10, D88314; Rat-mCpn10, P26772; Human-mCpn10, X75821; Saccharomyces-mCpn10, P38910. With respect to chl-Cpn10 and Cpn20, following regions were used for an alignment: chl-Cpn10, residues 50-139; Spinach-Cpn20-N, 62-157; Spinach-Cpn20-C, 160-255; Arabidopsis-Cpn20-N, 60-154; Spinach-Cpn20-C, 158-253.

Figure 3. Time-course of folding of citrate synthase in the presence of His-tagged co-chaperonins.

(a) The total extract of *E. coli* expressing His-tagged chl-Cpn10 (T) and the preparation used for the experiment (E) were subjected to SDS-PAGE. The CBB stained gel is shown. (b) Unfolded citrate synthase (CS) was diluted to a concentration of 150 nM with a solution containing $GroEL_{14}$ at 0ÚC. The temperature was immediately adjusted to 35ÚC and both 2 mM ATP and the following additions were made: closed squares, 300 nM $GroES_7$; rhombi, 300 nM His-tagged $Cpn20_4$; open triangles, 600 nM His-tagged $Cpn10_7$; closed triangles, 300 nM His-tagged $Cpn10_7$; circles, none.

Figure 4. Immunogold labeling of the transgenic Arabidopsis overproducing chl-Cpn10.

(a) Transgenic Arabidopsis overproducing chl-Cpn10 was prepared. The extracts of leaves of transgenic and wild-type Arabidopsis were subjected to SDS-PAGE and

subsequent immunoblotting.

(b, c, d) The transgenic Arabidopsis seedlings were germinated on agar medium in the dark. 4-day dark-grown and 1-day light–grown (b, c) and 7-day dark-grown (d) cotyledons were used for immunogold analysis. The left chloroplast shown in (b) is enlarged in (c). Immunogold labeling was performed with the antiserum against the His-tagged chl-Cpn10. Bar, 500 μm: C, chloroplast; E, etioplast, M, mitochondrion; V, vacuole; EM, extracellular matrix.

Figure 5. Immunoblot analysis of the membrane and the soluble fractions of chloroplasts for chl-Cpn10 and Cpn20.

The chloroplasts were disrupted by osmotic shock and separated by centrifugation. Total chloroplast protein (lane T), the membrane fraction (lane P) and the soluble fraction (lane S) that contained 3 μg, 10 μg and 21 μg of chlorophyll, respectively, were subjected to immunoblotting. Prestained size markers are given at the left.

Figure 6. Accumulation of mRNA for chl-Cpn10 in Arabidopsis.

(a) The total RNAs from leaves, stems and roots were subjected to Northern blotting probed with a radiolabeled chl-Cpn10 cDNA.

(b, c) Arabidopsis plants grown on soil were transferred from a normal growth temperature of 22ÚC to 35ÚC (heat treatment). The total RNAs from heat-treated leaves were subjected to Northern blotting (b) and quantitative RT-PCR (c). The heat-treatment time is shown.

Figure 7. Accumulation of chl-Cpn10 and Cpn20 in Arabidopsis seedlings during germination and greening.

Total protein (20 μg) extracted from seedlings grown under various conditions were subjected to SDS-PAGE and subsequent immunoblotting, respectively. Immunoblotting with the antiserum against the His-tagged chl-Cpn10 is shown (top). The antiserum against Arabidopsis Cpn20 was used as a control (middle). The number above each lane represents the day after germination in the dark: "4D1L" means that seedlings grown in

the dark for 4 days (4D) were transferred and kept in continuous light for a day (1L). The blotted membrane was stained with CBB (bottom).

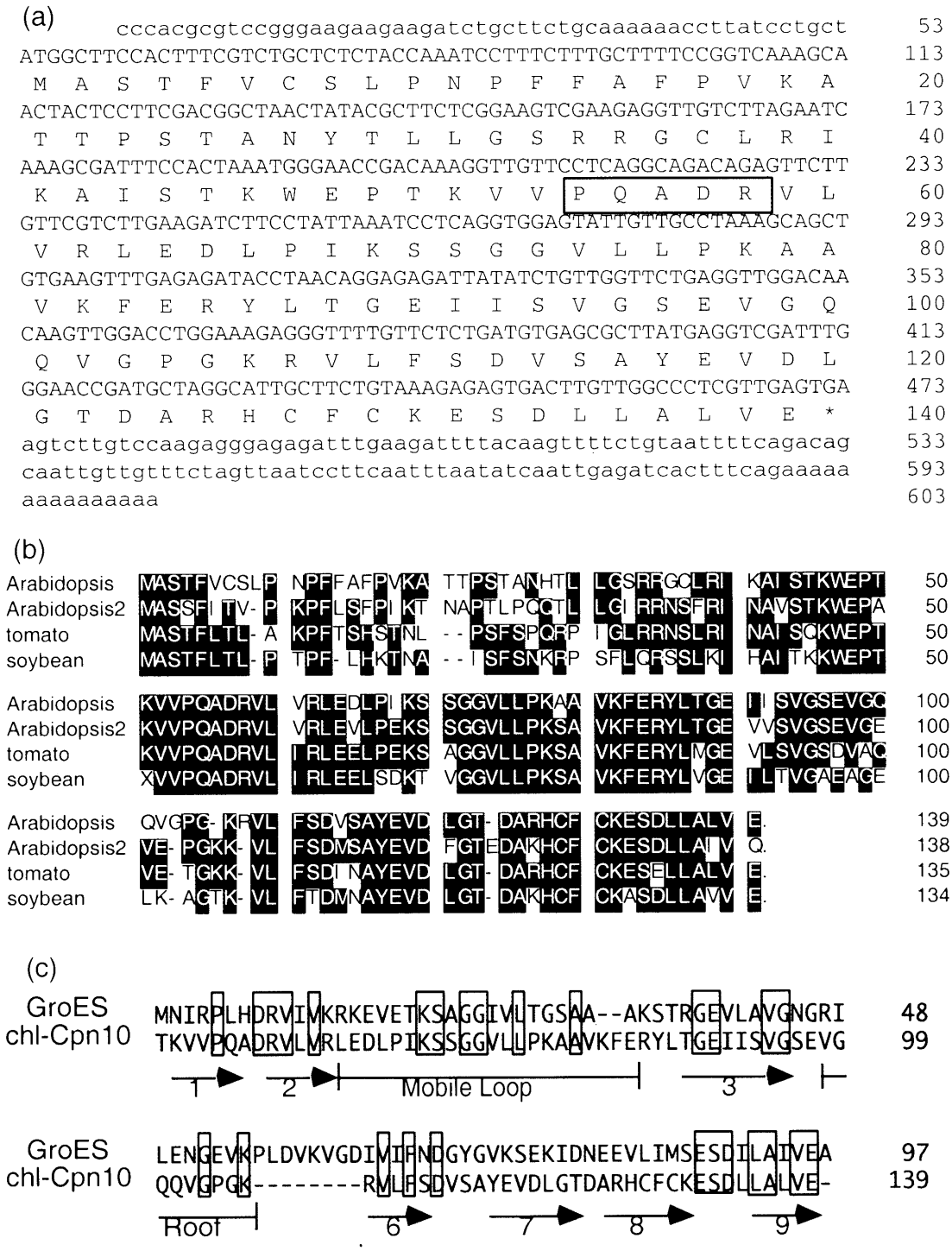


Figure 1 Koumoto *et al.*

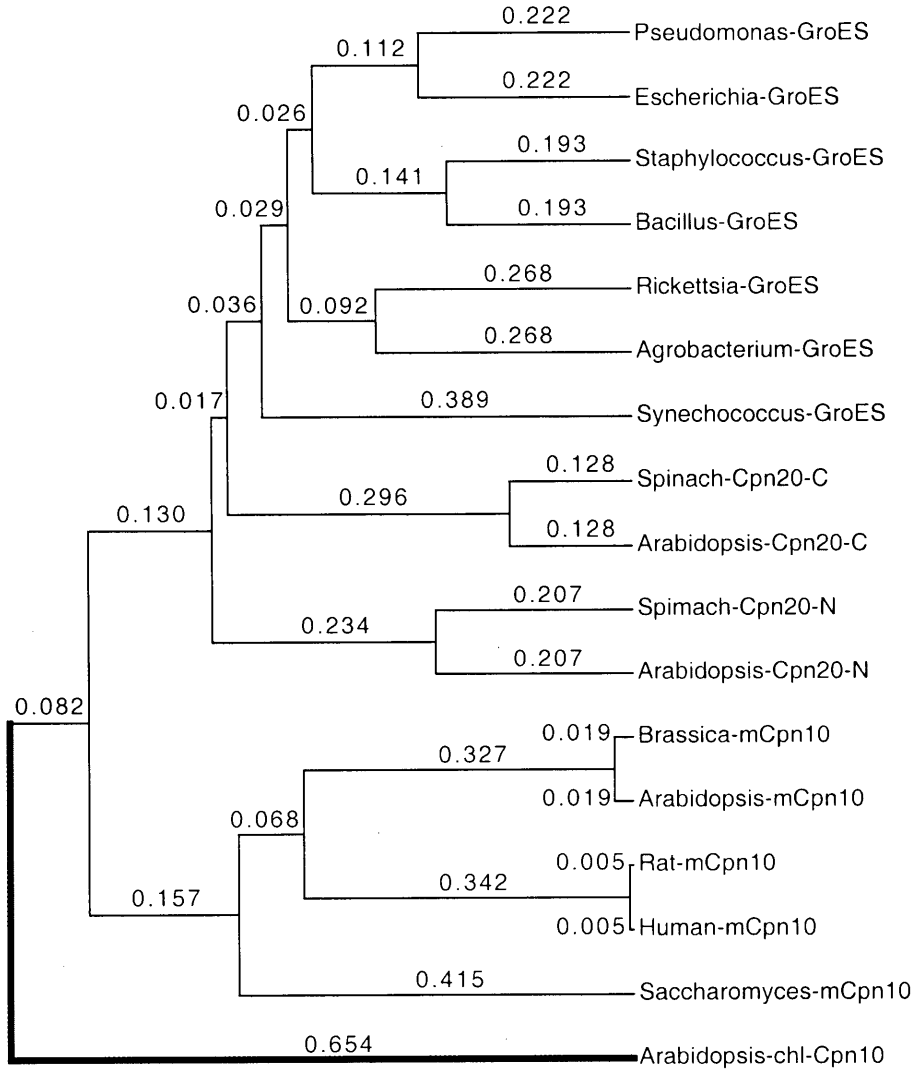


Figure 2 Koumoto *et al.*

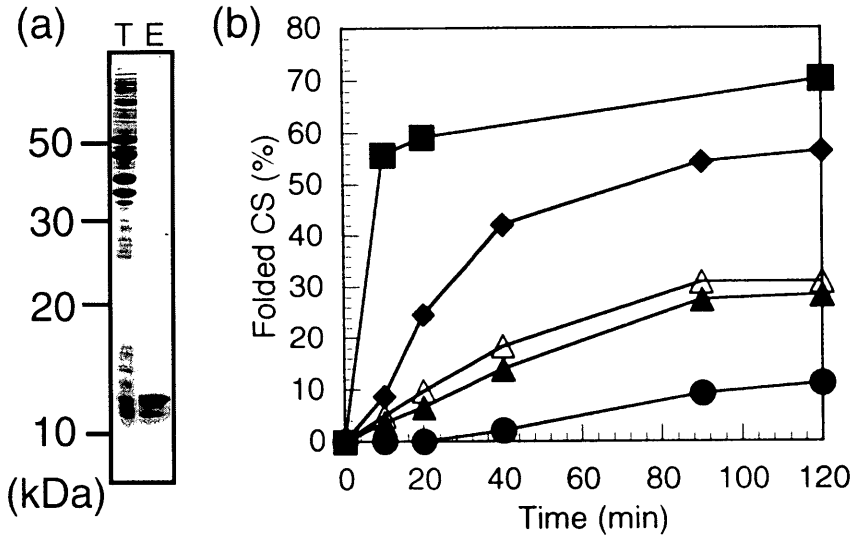


Figure 3 Koumoto *et al.*

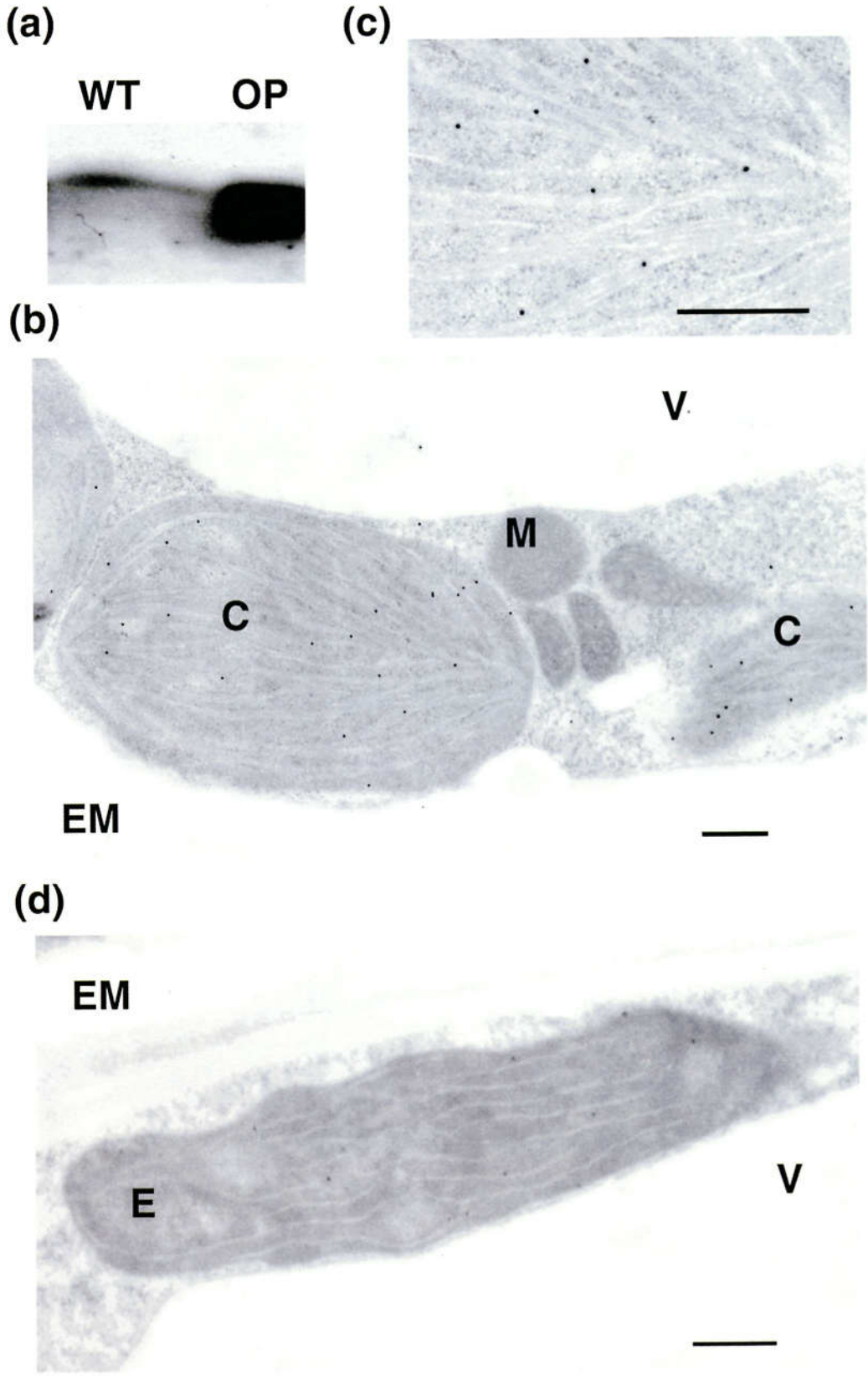


Figure 4 Koumoto *et al.*

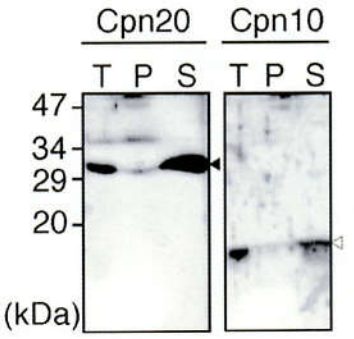


Figure 5 Koumoto *et al.*

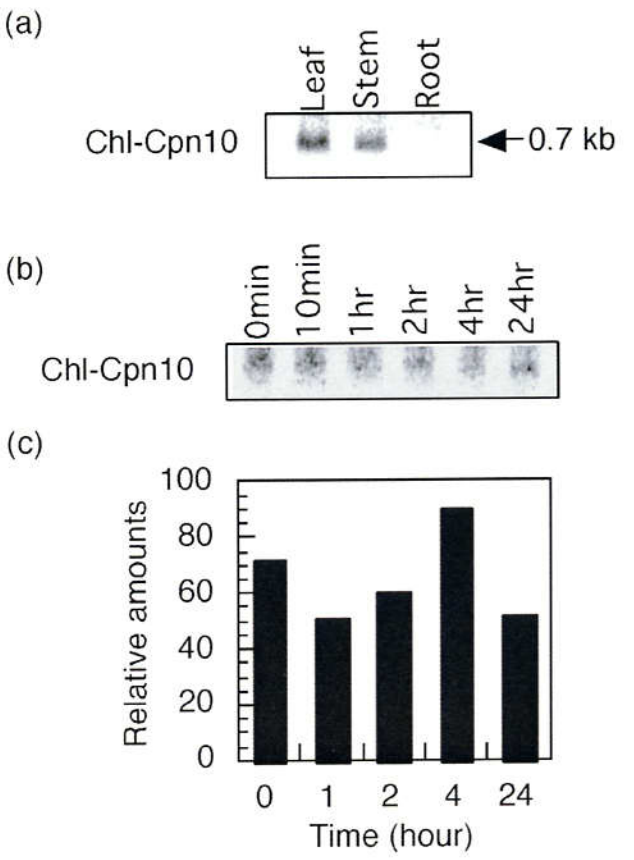


Figure 6 Koumoto *et al.*

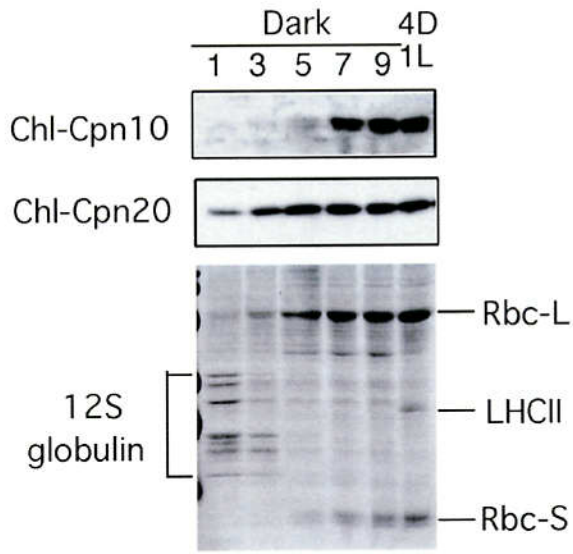


Figure 7 Koumoto *et al.*

参考論文目録

Sawano, H., Koumoto, Y., Ohta, K., Sasaki, Y., Segawa, S. and Tachibana, H. (1992) Efficient *in vitro* folding of the three-disulfide variants of hen lysozyme in the presence of glycerol. *FEBS Lett.* **303**, 11-14

Tachibana, H., Ohta, K., Sawano, H., Koumoto, Y. and Segawa, S. (1994) Relationship between the optimal temperature for oxidative refolding and the thermal stability of refolded state of hen lysozyme three-disulfide derivatives. *Biochemistry* **33**, 15008-15016.

Yokota, A., Izutani, K., Takai, M., Kubo, Y., Noda, Y., Koumoto, Y., Tachibana, H and Segawa, S. (2000) The transition state in the folding-unfolding reaction of four species of three-disulfide variant of hen lysozyme: the role of each disulfide bridge. *J. Mol. Biol.* **295**, 1275

Volume 303, number 1, 11–14 FEBS 11018 May 1992

Efficient in vitro folding of the three-disulfide derivatives of hen lysozyme in the presence of glycerol

Hiroshi Sawano[a], Yasuko Koumoto[b], Katsuyuki Ohta[b], Yukio Sasaki[a], Shin-ichi Segawa[c] and Hideki Tachibana[a,b]

[a]*Graduate School of Science and Technology and* [b]*Department of Biology, Faculty of Science, Kobe University, Rokkodai, Nada-ku, Kobe 657, Japan* and [c]*Department of Physics, School of Science, Kwansei Gakuin University, Nishinomiya 662, Japan*

Received 3 March 1992

Four derivatives of hen lysozyme, each lacking one native disulfide bond of the four in authentic lysozyme, were produced in *Escherichia coli* by expressing synthetic mutant genes. In the reoxidation reaction of the reduced derivatives purified from inclusion bodies, the addition of glycerol significantly enhanced the efficiency of folding and 'correct' disulfide bond formation. This enabled simple chromatographical purification of refolded materials. Purified 3SS-derivatives all showed lytic activities and secondary structures comparable to authentic lysozyme, which directly showed that none of the four native disulfide bonds is a prerequisite for 'correct' in vitro folding.

Protein folding; Disulfide bond; Lysozyme; Renaturation; Glycerol

1. INTRODUCTION

The contribution of individual disulfide bonds to the folding and stability of proteins which have multiple disulfide bonds has only been examined in a few proteins [1–5]. The detailed pathway of folding has been presented only for BPTI [1]. We have synthesized, and expressed in *E. coli*, the genes for all the four 3SS-derivatives of hen lysozymes, $\Delta 1$, $\Delta 2$, $\Delta 3$ and $\Delta 4$, in which one of the four native disulfide bonds, Cys^{6}–Cys^{127}, Cys^{30}–Cys^{115}, Cys^{64}–Cys^{80} and Cys^{76}–Cys^{94} (hereafter referred to as disulfide bonds 1, 2, 3 and 4, respectively), was opened by substitution of Ser for Cys residues (Fig. 1a,b). We show efficient conditions for the folding and 'correct' disulfide bond formation of the reduced 3SS-derivatives, and the enzymatic acitivity and secondary structure of the refolded 3SS-derivatives.

2. MATERIALS AND METHODS

2.1. *Materials*

A direct expression vector, pYK1, has been described [6]. *E. coli* strain AD18 (*Δ (lac-proAB), lon*-100, *tsx*::Tn*5*/F'[*lacI*^{q}, *lacZΔM*15, *lacY*$^{+}$, *proA*$^{+}$, *proB*$^{+}$]) was used for expression.

2.2. *DNA synthesis*

Oligodeoxyribonucleotides were synthsized on an Applied Biosystems 381A DNA synthesizer at the Research Center for Molecular Biology, Kobe University. Two sets, coding and non-coding, of 14 oligomers constitute the total gene (Fig. 1c). For two sets of 8 oligomers among them which contained the codon or anticodon sequence for a Cys residue, oligomers with the codon (or anticodon) sequence of Ser substituted for Cys were also synthesized. Manipulations of DNAs were carried out as described [7]. The 5'-end of the gene had a *Hin*dIII cohesive sequence. It also had a *Dra*I site (TTT'AAA) incorporated to facilitate excision of the gene from the cloning plasmid with the codon AAA, coding for the N-terminal lysine residue, exposed as a blunt end for its insertion to pYK1. The 3'-end contained a TAA termination codon and a *Bam*HI cohesive sequence.

2.3. *Expression of mutant genes and purification of the polypeptides*

Expression was carried out as described [8] except that we used *E. coli* strain AD18 and LB-medium containing ampicillin (25 μg/ml) and kanamycin (50 μg/ml). Harvested cells were disrupted by sonication and inclusion bodies were prepared as described [9]. They were solubilized in 8 M urea and 50 mM DTT, and reduced 3SS-derivatives were purified with cation-exchange (Mono-S) and gel-permeation (Sephadex G-75) chromatographies, freeze-dried and stored frozen under nitrogen.

2.4. *Peptide analysis*

Reduced and carboxamidomethylated [10] protein in 100 mM Tris-HCl, pH 8.0, was digested with TPCK-trypsin (E:S=1:100 by weight) at 30°C for 3 h or more. Tryptic peptides were separated with reversed-phase HPLC. The amino acid composition of each peptide was determined as described [11].

2.5. *Reoxidation of reduced protein*

Reoxidation and formation of disulfide bonds were carried out essentially as described [12], in 100 mM Tris-acetate, 1 mM EDTA, pH 7.8, 6 mM GSH and 0.6 mM GSSG, at the protein concentration of 3.3 μM, with the modification that the indicated amounts of glycerol were added and the indicated temperature was used. After 2 h of reoxidation the remaining (if any) thiol groups were carboxamidomethylated, and the reoxidation reaction was terminated by acidifying to below pH 5. The oxidized protein was purified by RPHPLC on a TSK TMS-250 column (4.6 mm × 7.5 cm; Tosoh) with a linear gradient of acetonitrile from 5 to 35% in 0.05% TFA.

Abbreviations: 3SS-, three disulfide bond-.

Correspondence address: H. Tachibana, Department of Biology, Faculty of Science, Kobe University, Rokkodai, Nada-ku, Kobe 657, Japan. Fax: (81) (78) 881-7593.

Published by Elsevier Science Publishers B.V.

2.6. *Other methods*

Lytic activity against *Micrococcus luteus* cells was measured as described [12]. Concentrations of the 3SS-derivatives were estimated as described [13]. CD spectra were measured using a J-600 spectropolarimeter (Japan Spectroscopic Co.) equipped with a thermostatically controlled cell holder.

3. RESULTS

We employed total gene synthesis because we also planned to synthesize 2SS- and 1SS-derivatives as well as fragments of hen lysozyme for future studies. The genes (Fig. 1) were synthesized in two steps of ligation. The recovered full-length gene was ligated to *Hin*dIII- and *Bam*HI-digested pUC18, with which *E. coli* JM109 was transformed, and the genes for Δ1 through Δ4 were cloned. Their nucleotide sequences were confirmed. Each gene was then excised by digestion with *Dra*I and *Bam*HI, gel-purified, and ligated to pYK1 which had been digested with *Nco*I, filled-in and digested with *Bam*HI. The expressed products were found mostly in inclusion bodies. The yield of the purified polypeptides were 1.0 to 6.2 mg per liter of culture. Tryptic peptide mapping of the reduced and carboxamidomethylated polypeptides confirmed correct substitution of Ser for Cys residues in all the four derivatives.

Reoxidation and formation of disulfide bonds under the conditions previously described for authentic lysozyme [12] gave materials with a low lytic activity (Table I) and a broad elution profile on RPHPLC (Fig. 2), probably due to a low stability of the folded state of these derivatives. When we lowered the temperature for the reoxidation reaction to 15°C, the activity values for the reoxidized materials increased and their elution profiles became sharp (not shown). Further decrease in reoxidation temperature, however, gave a lower activity and a broad profile again. Next, we added glycerol, which has been known to increase the stability of the native state of proteins [14] to the reoxidation solution, and found that the elution profile of the reoxidized materials became very sharp (Fig. 2). The elution time of the main peak for Δ1 through Δ4 nearly coincided with that for authentic lysozyme. Rechromatography of the main-peak fraction with RPHPLC or cation-exchange HPLC gave an apparently single peak for all the four derivatives (not shown). The tryptic-peptide maps for the reoxidized and purified derivatives, and the results of the amino acid analysis for the peptide peaks which newly appeared on the reoxidation reaction, were consistent with the formation of the three 'correct' disulfide bonds in each of the four derivatives (not shown).

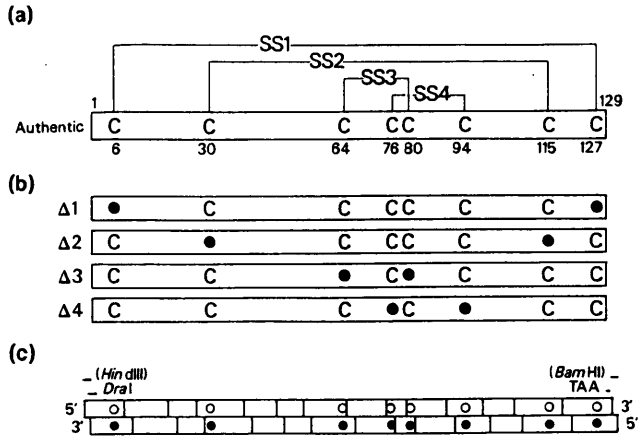


Fig. 1. (a) Eight cysteine residues ('C') and four disulfide bridges (SS1 to SS4) in hen lysozyme. (b) The four 3SS-derivatives. Filled circles represent substitutions of Ser for Cys residues. (c) Nucleotide regions covered by the 28 oligomers (rectangular blocks) which make up a synthetic gene. Open and filled circles represent the codon and anticodon sequences for Cys residues.

The activity of the purified 3SS-derivatives was comparable to that of authentic lysozyme (Table I). Circular dichroism spectra (Fig. 3) showed that they had secondary structures comparable to that of authentic lysozyme. The mean residue ellipticity values were independent of the protein concentration (from 4 to 20 μM) indicating that the observed secondary structure was intramolecularly formed. Their spectral shapes were different from each other. In particular, Δ2 showed a marked difference from the rest.

4. DISCUSSION

We have shown that none of the four disulfide bridges of native lysozyme is obligatory for correct in vitro

Table I

Lytic activities of the 3SS-derivatives of lysozyme

3SS-species	Recombinant synthesized in *E. coli*[a]		Partial oxidation or reduction	Recombinant secreted from yeast [16]
	Refolded at 37°C, no glycerol[b]	Refolded at 15°C, in 20% glycerol[c]		
Δ1	9 (%)[d]	68 ± 2[e] (%)[d]	40–50 [15], 58 [18] (%)[d]	ND[f] (%)[g]
Δ2	10 (%)[d]	76 ± 7 (%)[d]	ND	28
Δ3	13 (%)[d]	88 ± 4 (%)[d]	40–50 [15]	23
Δ4	19 (%)[d]	142 ± 11 (%)[d]	40–50 [15]	94

[a]This study, [b]reoxidized materials, not purified, [c]purified main peak fraction, [d]values relative to authentic hen lysozyme, [e]mean and SD for three measurements, [f]not determined, [g]values relative to authentic human lysozyme.

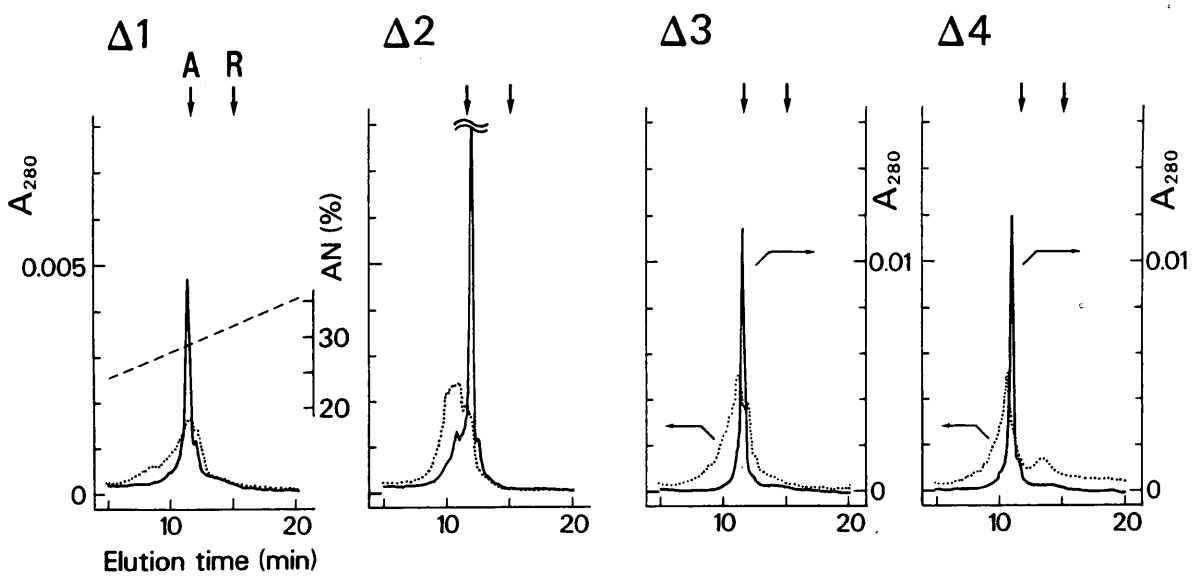


Fig. 2. Elution profiles of the 3SS-derivatives reoxidized at 37°C without glycerol (dotted lines) and those at 15°C with 20% glycerol (full lines). The gradients in acetonitrile are in broken lines. The amount of protein reoxidized was about 4 μg, the filtrate of which was subjected to RPHPLC. A and R indicate the elution positions of authentic and reduced lysozymes, respectively.

folding of hen lysozyme to enzymatically active forms. The same conclusion has been drawn [4] using the materials obtained through reoxidation of reduced and partially, randomly, carboxymethylated hen lysozyme. Our approach was more direct, and since all the four 3SS-derivatives were obtained in significant amounts due to efficient folding in the presence of glycerol, we could also examine their secondary structures, which turned out not to be identical. Previously, except for the 3SS-derivative corresponding to Δ1, it was difficult to obtain other 3SS-derivatives in sufficient amounts. In particular, the derivative corresponding to Δ2 could not be obtained [15]. In recent studies on human lysozyme, although the derivative corresponding to Δ4 was synthesized in large amounts, the derivative corresponding to Δ1 was not secreted in the yeast expression system, and those corresponding to Δ2 and Δ3 were secreted inefficiently [16,17]. Detailed analyses of the structural differences among the 3SS-derivatives constructed here will help to understand the interactions inducing and stabilizing intramolecular structures in the folding of hen lysozyme.

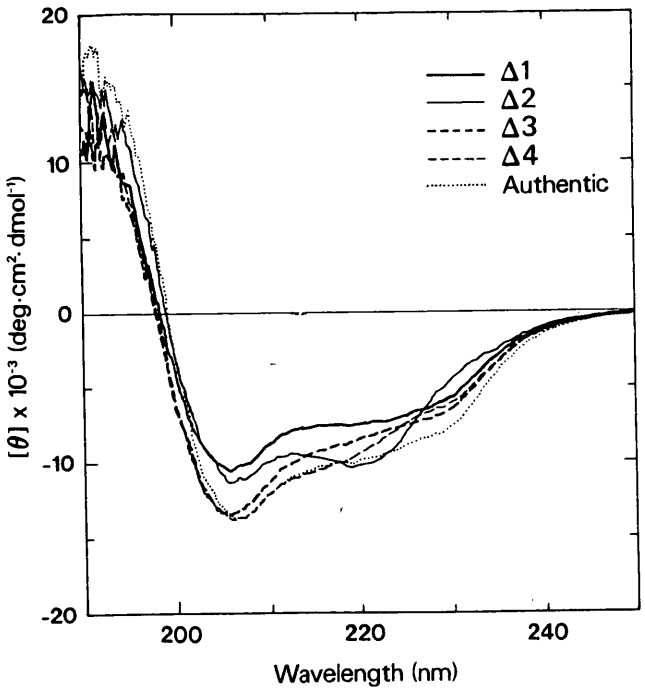


Fig. 3. Circular dichroism spectra for purified Δ1 (thick line), Δ2 (thin line), Δ3 (thick broken line), Δ4 (thin broken line) and authentic lysozyme (dotted line) in 20 mM phosphate adjusted to pH 3.9 with sodium hydroxide and at 25°C. Protein concentrations were 4.0, 3.4, 3.6, 3.7 and 3.5 μM, respectively.

Acknowledgements: We thank Dr. T. Yura, Institute for Virus Research, Kyoto University, for providing AD18, Drs. M. Kitakawa, K. Isono, and Y. Fukami for advice and facilities on DNA synthesis, and Prof. Y. Nishizuka for encouragement and support. This work was partly supported by grants-in-aid for scientific research from the Ministry of Education, Science and Culture, Japan.

REFERENCES

[1] Creighton, T.E. (1975) J. Mol. Biol. 95, 167–199.
[2] States, D.J., Creighton, T.E., Dobson, C.M. and Karplus, M. (1987) J. Mol. Biol. 195, 731–739.
[3] Marks, C.B., Naderi, H., Kosen, P.A., Kuntz, I.D. and Anderson, S. (1987) Science 235, 1370–1373.
[4] Acharya, A.S. and Taniuchi, H. (1977) Proc. Natl. Acad. Sci. USA 74, 2362–2366.
[5] Anderson, W.L. and Wetlaufer, D.B. (1976) J. Biol. Chem. 251, 3147–3135.
[6] Tachibana, H., Sasaki, Y., Sawano, H. and Kitikawa, M. (1990) Protein Eng. 3, 371.

[7] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning, A laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, New York.
[8] Miki, T., Yasukochi, T., Nagatani, H., Furuno, M., Orita, T., Yamada, H., Imoto, T. and Horiuchi, T. (1987) Protein Eng. 1, 327–332.
[9] Nagai, K. and Thøgersen, C. (1987) Methods Enzymol. 153, 461–481.
[10] Hirs, C.H.W. (1967) Methods Enzymol. 11, 199–203.
[11] Heinrikson, R.L. and Meredith, S.C. (1984) Anal. Biochem. 136, 65–74.
[12] Saxena, V.P. and Wetlaufer, D.B. (1970) Biochemistry 9, 5015–5023.
[13] Gill, S.C. and von Hippel, P.H. (1989) Anal. Biochemistry 182, 319–326.
[14] Gekko, K. and Timasheff, S.N. (1981) Biochemistry 20, 4677–4686.
[15] Acharya, A.S. and Taniuchi, H. (1976) J. Biol. Chem. 251, 6934–6946.
[16] Taniyama, Y., Yamamoto, Y., Nakao, M., Kikuchi, M. and Ikehara, M. (1988) Biochem. Biophys. Res. Commun. 152, 962–967.
[17] Inaka, K., Taniyama, Y., Kikuchi, M., Morikawa, K. and Matsushima, M. (1991) J. Biol. Chem. 266, 12599–12603.
[18] Radford, S.E., Woolfson, D.N., Martin, S.R., Lowe, G. and Dobson, C.M. (1991) Biochem. J. 273, 211–217.

Reprinted from Biochemistry, 1994, 33.

Relationship between the Optimal Temperature for Oxidative Refolding and the Thermal Stability of Refolded State of Hen Lysozyme Three-Disulfide Derivatives†

Hideki Tachibana,*,‡,§ Katsuyuki Ohta,‡,⊥ Hiroshi Sawano,§ Yasuko Koumoto,∥,# and Shin-ichi Segawa∥

Department of Biology, Faculty of Science, and Graduate School of Science and Technology, Kobe University, Rokkodai, Nada-ku, Kobe 657, Japan, and Department of Physics, School of Science, Kwansei Gakuin University, Nishinomiya 662, Japan

Received March 31, 1994; Revised Manuscript Received October 10, 1994⊗

ABSTRACT: The temperature dependence of the efficiency of oxidative refolding was examined for hen lysozyme three-disulfide derivatives produced in *Escherichia coli*. Each derivative was designed to lack one of the four disulfide bridges in authentic lysozyme: Δ1 (Cys6→Ser, Cys127→Ser), Δ2 (Cys30→Ser, Cys115→Ser), Δ3 (Cys64→Ser, Cys80→Ser), Δ4 (Cys76→Ser, Cys94→Ser), Δ2Ala (Cys30→Ala, Cys115→Ala), and Δ4Ala (Cys76→Ala, Cys94→Ala). The optimal refolding temperature was lowest for Δ1 (19 °C) and highest for Δ4A1a (30 °C). The chromatographically purified, completely refolded three-disulfide species were not stable above the optimal refolding temperature in the presence of glutathione. The stability of each of them was determined from the far-UV CD thermal denaturation measurement at pH 3.9 in the absence of glutathione, where the denaturation was reversible. The transition temperature was lowest for Δ1 and highest for Δ4Ala. Precise values of difference in the transition temperature among the three-disulfide derivatives were found to correlate with those in the optimal refolding temperature. Next, the effect of glycerol, which has been shown to increase the refolding efficiency [Sawano et al. (1992) *FEBS Lett. 303*, 11−14], was examined for Δ1 in detail. The optimal temperature for refolding increased by 3−4 °C with the increase in glycerol concentration by 10%. The amount of increase in the optimal refolding temperature was nearly equal to the amount of the increase in thermal stability in the presence of glycerol of refolded and purified Δ1. Taken together, there exists a parallel relationship for the three-disulfide derivatives between the optimal refolding temperature and the thermal stability of the correctly refolded state. The observation provides the basis for the optimization of the refolding temperature of engineered proteins of low stability.

Expression of recombinant, foreign genes in *Escherichia coli* often leads to the formation of inclusion bodies in which recombinant proteins are produced as insoluble aggregates (Mitraki & King, 1989). The polypeptides solubilized and purified from inclusion bodies under denaturing and reducing conditions should be renatured into functional proteins of defined tertiary structure. It is therefore needed to develop efficient methods of renaturation. For the proteins which contain disulfide bridges, the refolding conditions which utilize reduced and oxidized forms of glutathione (Saxena & Wetlaufer, 1970) are widely used. The detailed reaction conditions, however, such as temperature, pH, salt concentration, and redox potential have to be optimized empirically.

Hen lysozyme contains four disulfide bridges. In spite of a number of studies on the folding reaction of lysozyme, it has not fully been clarified how these bridges are involved in the reaction process (Ristow & Wetlaufer, 1973; Anderson & Wetlaufer, 1976; Acharya & Taniuchi, 1976, 1977, 1982; Taniyama et al., 1988; Radford et al., 1991). To study the role of each bridge in the folding as well as in the stabilization of this protein, we have undertaken the construction of all the four molecular species of hen lysozyme three-disulfide derivatives in which Cys residues that form a disulfide bridge in authentic lysozyme were replaced by Ser: Δ1[1] (Cys6→Ser, Cys 127→Ser), Δ2 (Cys30→Ser, Cys115→Ser), Δ3 (Cys64→Ser, Cys80→Ser), and Δ4 (Cys76→Ser, Cys94→Ser) (Sawano et al., 1992). The 3SS derivatives purified from inclusion bodies under denaturing and reducing conditions did not efficiently refold under the conditions which have been reported by Saxena and Wetlaufer (1970) to be optimal for the oxidative refolding (i.e., regeneration of correct disulfide bridges and of nearly full enzymatic activity) of authentic lysozyme. We showed, however, that they refolded at low temperatures and/or in the presence of glycerol.

In this study, to examine the factors which influence the refolding efficiency, we determined the optimal refolding temperatures for the four 3SS derivatives as well as two

† This work was partly supported by grants-in-aid for scientific research from the Ministry of Education, Science and Culture, Japan, and by a grant to S.S. from Hyogo Science and Technology Association.

* Author to whom correspondence should be addressed.

‡ Department of Biology, Kobe University.

§ Graduate School of Science and Technology, Kobe University.

⊥ Present address: Chemical Process Research and Development, Production Department, Shionogi & Co. Ltd., 1-3, Kuise Terajima 2-chome, Amagasaki 660, Japan.

∥ Department of Physics, Kwansei Gakuin University.

Present address: Department of Cell Biology, National Institute for Basic Biology, Okazaki 444, Japan.

⊗ Abstract published in *Advance ACS Abstracts*, November 15, 1994.

[1] Abbreviations: 3SS, three-disulfide; Δ1, a 3SS derivative of hen lysozyme in which Cys residues 6 and 127 are replaced by Ser residues; (for other 3SS derivatives, the amino acid substitutions are as follows: Δ2 (Cys30→Ser, Cys115→Ser), Δ3 (Cys64→Ser, Cys80→Ser), Δ4 (Cys76→Ser, Cys94→Ser), Δ2A1a (Cys30→Ala, Cys115→Ala), and Δ4Ala (Cys76→Ala, Cys94→Ala)); T_{opt}, optimal temperature for the oxidative refolding from reduced state; T_m, midpoint temperature for the unfolding reaction (without the opening of disulfide bridges); RPHPLC, reversed-phase high-performance liquid chromatography; GSH, reduced glutathione; GSSG, oxidized glutathione; RNase A, bovine pancreatic ribonuclease A; α-LA, α-lactalbumin.

0006-2960/94/0433-15008$04.50/0

Cys→Ala 3SS derivatives: Δ2Ala (Cys30→Ala, Cys115→Ala) and Δ4Ala (Cys76→Ala, Cys94→Ala). The optimal temperatures were related with the thermal stability of the derivatives correctly refolded.

MATERIALS AND METHODS

Hen Lysozyme 3SS Derivatives. Construction of the synthetic genes for Δ1−Δ4 and their direct expression in *E. coli*, based on the reported method (Miki et al., 1987), have been described (Sawano et al., 1992). The genes for Δ2Ala and Δ4Ala were constructed by using the synthetic oligonucleotides, which contained the codons for Ala residues in place of the respective Cys residues. The nucleotide sequence of each recombinant gene cloned in plasmid was confirmed. The genes were then incorporated to the direct expression vector pYK1 (Tachibana et al., 1990), and expressed in a similar way as for Δ1−Δ4.

Purification of the polypeptides under reducing (and denaturing) conditions was carried out as follows. Harvested cells (from 5.4-L culture) were disrupted by sonication in 200 mL of 50 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 0.1 mM phenylmethanesulfonyl fluoride. Inclusion bodies were prepared as described (Nagai et al., 1987). They were then solubilized in 24 mL of 8 M urea, 10 mM sodium phosphate, 1 mM EDTA, and 50 mM DTT, pH 7.9, at 37 °C for 2 h. After the pH was lowered to 5.0 with acetic acid, the solution was filtered through a 0.8-μm filter and subjected to cation-exchange chromatography on a SE column (Productive, bps Separations Ltd.) which had been equilibrated with 8 M urea, 10 mM sodium acetate, 1 mM EDTA, and 10 mM DTT, pH 5.0, and eluted with step gradients in NaCl concentration. Fractions containing polypeptides of 14 kDa, as monitored with SDS−PAGE, were subjected to gel permeation chromatography on a Sephadex G-75 column (2.5 × 50 cm), and reduced polypeptides were eluted with 0.1 M acetic acid, freeze-dried, and stored frozen under nitrogen. The concentration of the purified, reduced 3SS derivatives were estimated by using $A_{280} = 2.64$ for 1 mg/mL protein, which was deduced based on the content of tryptophan and tyrosine residues (Gill & von Hippel, 1989). The yield of purified polypeptide was 7−10 mg/L of culture.

The substitutions of Ser or Ala for Cys residues in the purified, reduced 3SS derivatives were confirmed by tryptic peptide mapping. Briefly, the polypeptides were carboxamidomethylated (Hirs, 1967) and digested with *N*-tosyl-L-phenylalanyl chloromethyl ketone-treated trypsin (Seikagaku Kogyo) (E:S = 1:100 in weight) in 100 mM Tris-HCl, pH 8.0, at 30 °C for 3 h. The tryptic peptides were separated with reversed-phase HPLC on TSK-ODS120T column (7.8 mm × 30 cm; Tosoh) with a 50-min gradient of acetonitrile from 1 to 50% in 0.1% TFA at 40 °C. The polypeptides in each peak were hydrolyzed with 6 N HCl containing 0.01% phenol, and the amino acids were derivatized with phenyl isothiocyanate (sequencing-grade, Wako Pure Chemical Industries, Ltd.) as described by Heinrikson and Meredith (1984). Phenylthiocarbamyl amino acids were separated and quantitated with reversed-phase chromatography on a TSK-ODS80TM column (4.6 mm × 25 cm; Tosoh) with a 20-min gradient from 3% acetonitrile, 50 mM sodium phosphate (pH 7.0), and 100 mM sodium perchlorate to 33% acetonitrile, 10 mM sodium phosphate (pH 7.0), and 20 mM sodium perchlorate.

Reoxidation of Reduced Protein and Purification of Refolded Protein. Reoxidation and formation of disulfide bonds were carried out in 100 mM Tris−acetate, 1 mM EDTA, pH 7.8, 6 mM GSH, and 0.6 mM GSSG as described (Saxena & Wetlaufer, 1970), except that various temperatures and, when specified, glycerol concentration (v/v %) were used. The protein concentration was 3.3 μM. For the measurement of the refolding efficiency at each temperature and glycerol concentration, 6−12 μg of protein was used. The temperature dependence of the pH of the refolding buffer was not corrected for. The shift in pH value due to the presence of glycerol was not more than 0.2. Before being mixed to start refolding, both the refolding buffer and the reduced protein in 0.1 N acetic acid were equilibrated at the reaction temperature. The refolding reaction was stopped by adding acetic acid to bring the solution pH below 5, and the solution was passed to activity measurement.

As the control in the reoxidation experiment, reduced authentic lysozyme was also prepared. Hen egg-white lysozyme (6× crystallized, Seikagaku Kogyo) was reduced in 0.2 M Tris-HCl, 1 mM EDTA, 8 M urea, pH 8.0, and 50-fold molar excess of DTT over the content of disulfide in the protein at 40 °C for 3 h; subjected to gel permeation chromatography in 0.1 N acetic acid on Sephadex G25; and freeze-dried.

To know the extent of conformational heterogeneity in the reoxidized proteins, the reoxidized protein solution was subjected to reversed-phase HPLC after being acidified below pH 5 or after the remaining, if any, thiol groups were carboxamidomethylated on a TSK TMS-250 column (4.6 mm × 7.5 cm, Tosoh) at 40 °C with a linear gradient of acetonitrile from 28 to 40% in 0.05% TFA. Native and reduced authentic lysozymes were used as references for elution position.

Reduced 3SS derivatives were reoxidized in a larger scale (300−900 μg) at the temperature and the glycerol concentration, which were determined to be optimal in the preceding experiments. The main peak fraction on the reversed-phase HPLC described above was recovered, freeze-dried, and used for stability measurement. The protein in the main peak fraction eluted as a single peak when rechromatographed on reversed-phase or cation-exchange (Asahipak ES-502C, 7.6 mm i.d. × 100 mm, Asahi Kasei Kogyo) HPLC.

Activity Measurement. The efficiency of refolding was expressed by the lytic activity observed for the reoxidized material. A portion of the solution that contained 3.3 μg of the reoxidized 3SS derivative protein was mixed with the solution of *Micrococcal luteus* cells (Seikagaku Kogyo) in 50 mM phosphate buffer, pH 6.20, which had been equilibrated at 25 °C, to the total volume of 1.10 mL and the final cellular concentration of 0.35 mg/mL. The time derivative of the 660-nm absorbance decreases 30 s after the mixing was taken as the lytic activity value and shown as a value relative to that for an equal amount of authentic lysozyme. When necessary, the reaction scale was increased for a better reproducibility. The relative activity values varied from one experiment to another by less than 4% (Δ1, Δ2, Δ3) or 7% (Δ2Ala, Δ4, Δ4Ala). (Here, 100% corresponds to the activity of authentic lysozyme.) The resultant deviation in each T_{opt} value was not more than ±2 °C. The presence of glycerol, GSSG, or GSH to the amount which was expected to be carried over to the assay solution did not decrease or

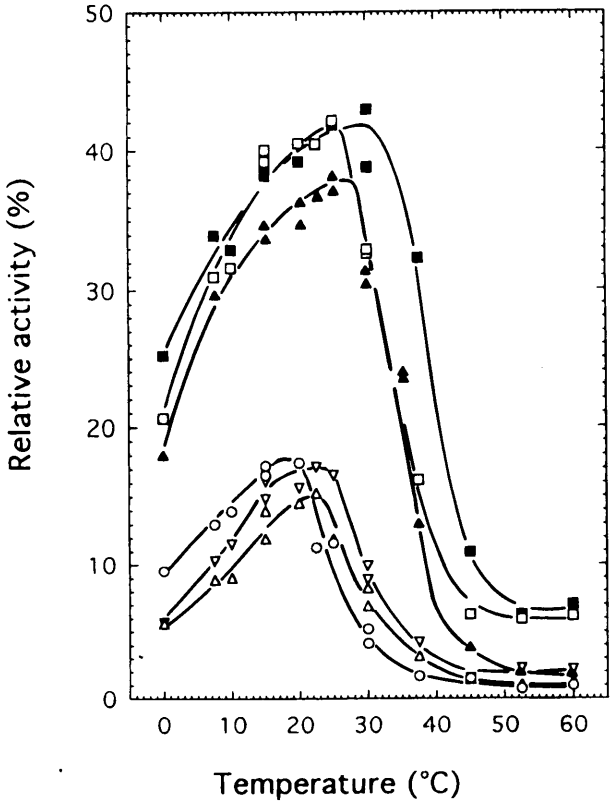


FIGURE 1: Lytic activity of the solution of each 3SS derivative refolded at various temperatures. Symbols and T_{opt} values are (○) Δ1, 19 °C; (△) Δ2, 22 °C; (▲) Δ2Ala, 27 °C; (▽) Δ3, 23.5 °C; (□) Δ4, 25 °C; (■) Δ4Ala, 30 °C.

increase in a control experiment the activity of authentic lysozyme.

Stability Measurement. The stability of chromatographically purified, completely refolded 3SS derivatives in the presence of reducing reagent and at a temperature higher than T_{opt} was examined as follows: The protein solution in 0.1 N acetic acid and the buffer containing Tris–acetate, EDTA, and reduced and oxidized glutathione both equilibrated at the temperature were mixed, resulting in the same solution conditions and protein concentration as used in the refolding experiment. After various intervals, the solution was acidified to pH < 5, filtered through a 0.2-μm filter, and analyzed with RPHPLC on a TSK TMS-250 column as described above. As a control, the buffer lacking the thiol reagents was also used.

CD thermal transition curves were obtained from the far-UV region CD spectra measured at various temperatures using a J-600 spectropolarimeter (Japan Spectroscopic Co.) equipped with a thermostatically controlled cell holder. The solvent was 20 mM phosphate, with or without glycerol, adjusted to pH 3.90 (at room temperature) with sodium hydroxide. Protein concentration was 3.0–6.4 μM. The temperature of the solution was directly measured using a thermistor thermometer.

RESULTS

Optimal Temperatures for the Refolding Reaction of 3SS Derivatives. The temperature dependence of the refolding reaction for the 3SS derivatives is shown in Figure 1. The ordinate shows the bacteriolytic activity of the reoxidized protein solution, in which a mixture of refolded proteins of various tertiary structure or disulfide pairing was expected to be contained. The temperature dependence for each derivative was qualitatively similar to the typical temperature profile of the enzyme activity: gradual increase to the maximal level and steep decrease with the increase in temperature. The optimal refolding temperatures were 19, 22, 27, 23.5, 25, and 30 °C for Δ1, Δ2, Δ2Ala, Δ3, Δ4, and Δ4Ala, respectively. Among the Cys→Ser derivatives, the result was Δ1 < Δ2 < Δ3 < Δ4 in the increasing order of T_{opt}. It was about 5 °C higher for the Cys→Ala derivative than for its counterpart Cys→Ser derivative. The observed bacteriolytic activity was much higher for Δ2Ala, Δ4, and Δ4Ala than for Δ1, Δ2, and Δ3. Under the same refolding conditions as above, T_{opt} for reduced authentic lysozyme was about 35 °C, and the relative bacteriolytic activity of its refolded solution amounted to 90% (not shown) in agreement with a previous paper (Saxena & Wetlaufer, 1970). These data were obtained in the reoxidation experiments of 50-min incubation time. Although kinetic experiments showed that in this time interval the reaction did not reach equilibrium for lower reaction temperatures, the experiments employing an incubation time of 200 min showed that qualitative as well as quantitative features described above were conserved: the newly estimated T_{opt}'s agreed with the original one within ±1 °C (not shown).

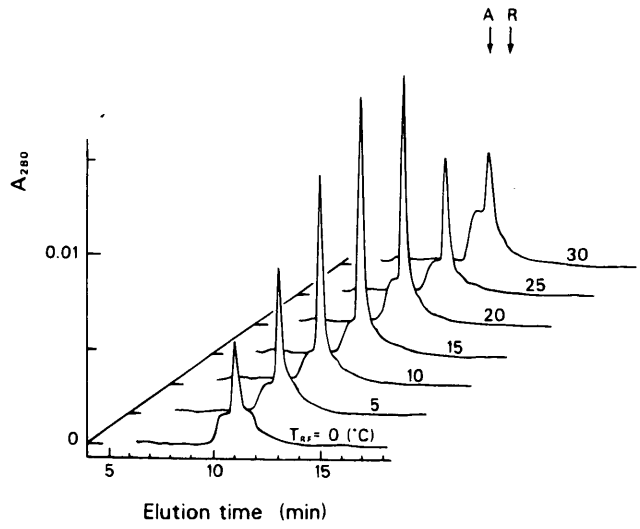


FIGURE 2: Elution profiles of the refolded Δ1 on RPHPLC. Each profile is labeled with the temperature, T_{RF}, at which the refolding reaction was carried out. 8.2 μg of reduced Δ1 was refolded, filtered through a 0.2-μm filter, and subjected to RPHPLC as described in Materials and Methods. Arrows A and R indicate the elution positions for authentic and reduced lysozymes, respectively.

To examine the conformational heterogeneity among the reoxidized proteins, the reoxidized solution was subjected to reversed-phase HPLC. The result for Δ1 is shown in Figure 2. A peak on a broad background was observed for each refolding condition of different temperature. The elution time for the peak was common among the profiles shown and nearly coincided with that of authentic lysozyme (position "A" in the figure). There existed a correlation between the height of the peak and the apparent lytic activity value shown for various temperatures in Figure 1. A similar result was obtained for other 3SS derivatives (not shown). Each temperature profile in Figure 1 therefore can be taken as to approximately represent the change with temperature in the fraction of correctly refolded 3SS derivative. (For a comparison of the fractions of correctly refolded molecule

between the profiles, however, it is necessary to take the specific activity of each 3SS derivative into consideration.) This point was supported by the following observations: First, the refolded 3SS derivatives isolated in a large scale from the peak fraction showed the far-UV CD spectra that indicated the presence of secondary structure of nearly the same amount as authentic lysozyme [the spectra of the Cys→Ser derivatives have been reported (Sawano et al., 1992); the spectra of Δ2Ala and Δ4Ala had the same characteristics as those for their respective, counterpart Cys→Ser derivatives (not shown)]. Second, the tryptic map of each of the purified, refolded derivatives was consistent with the formation of each set of the three, out of four, native disulfide bridges in them (not shown). Third, the purified refolded derivatives showed lytic activities which were comparable (60−140%) to authentic lysozyme (Sawano et al., 1992, for the Cys→Ser derivatives; the relative activities at 25 °C for Δ2Ala and Δ4Ala were 72 and 90%, respectively). Furthermore, when we calculated the expected lytic activity of the crude, refolded protein solution by using the estimated amount of the "correctly" refolded protein of the peak fraction in the chromatogram shown in Figure 2 and the activity value of the purified derivatives described above, it nearly agreed with the observed lytic activity value shown in Figure 1.

Thermal Stability of the Refolded and Purified 3SS Derivatives and Its Relationship with T_{opt}. The decrease in the refolding efficiency in the temperature range above T_{opt} is considered to be due to thermal instability of intermediate species and/or the completely refolded species in the folding pathway. We examined the stability of the completely refolded and chromatographically purified Δ1, which was subjected to the same solution conditions as those in the refolding experiment, by monitoring the RPHPLC elution profiles (Figure 3). A reaction temperature of 30 °C was chosen, where the recovered activity in the refolding experiment had been shown to decrease to about 25% of the optimal one (Figure 1). After 90-s incubation, the height of the peak for the completely and correctly refolded Δ1 decreased to 80%, and a broad background, which was indicative of disulfide-reshuffled proteins of nonnative structure, appeared. After 15 min, the peak height decreased to 29%, and after 50 min it decreased to 22%, with the elution profile becoming very similar to the one shown for a refolding temperature (T_{RF}) of 30 °C in Figure 2. The fraction of the intact Δ1 remaining after the 50-min incubation roughly agreed with the above-mentioned fraction of the recovered activity at 30 °C to the optimal one. In a control experiment in which the redox reagents were omitted, the sharp peak was completely conserved after 50-min incubation. Similar results were obtained for the other 3SS derivatives (not shown). The results indicate that the correctly refolded 3SS derivatives are instable at the temperature above T_{opt} in the present conditions used for refolding.

It was difficult with spectrophotometric methods to accurately determine the denaturation temperature of the completely refolded and chromatographically purified 3SS derivatives under the conditions used for refolding, i.e., in the presence of the redox reagents and at the dilute protein concentration. Even in the absence of the redox reagents, the solution of the refolded and purified derivatives showed aggregation at neutral pH regions. Therefore, we carried

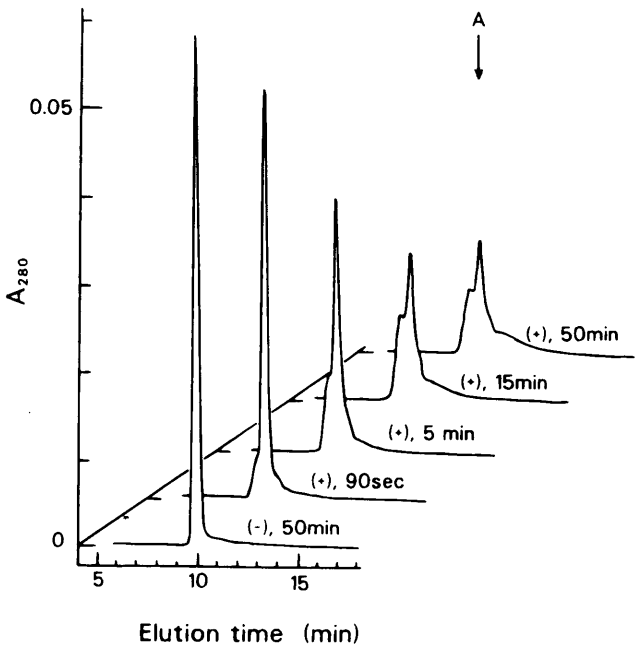


FIGURE 3: Stability of the completely refolded and chromatographically purified Δ1 under the refolding solution conditions and at 30 °C. The label to each RPHPLC elution profile indicates the presence (+) or absence (−) (as a control) of thiol reagents (6 mM GSH and 0.6 mM GSSG) and the incubation time. The amount of protein for each chromatogram was 7.5 μg (including a possible loss during filtration). Arrow A indicates the elution position for authentic lysozyme. It is not the same as the one in Figure 2 due to the aging of the column.

out far-UV CD thermal denaturation measurement at pH 3.9 without the redox reagents, although the denaturation temperature under this condition is expected to be higher than that under the refolding condition. At pH 3.9, reversible transition curves were obtained for all the derivatives (Figure 4). The transition midpoints (T_m) were 39.2, 42.8, 48.2, 45.2, 44.9, and 51.6 °C (±0.6 °C) for Δ1, Δ2, Δ2Ala, Δ3, Δ4, and Δ4Ala, respectively. Among the Cys→Ser derivatives, the result was Δ1 < Δ2 < Δ3 < Δ4 in the increasing order of T_m, and the Cys→Ala derivatives had the T_m which was higher than that of the counterpart Cys→Ser derivative by about 6 °C. Furthermore, the T_m data obtained in our preliminary differential scanning calorimetry measurements (not shown) agreed with these results.

T_{opt} and T_m for each 3SS derivative were plotted in Figure 5. A nearly parallel relationship between the two quantities was evident. The difference between them, which was almost constant among the derivatives, was 20.9 ± 0.7 °C.

Effect of Glycerol on T_{opt} and T_m of Δ1. It has been known that glycerol stabilizes the native state of protein through the mechanism of "preferential hydration" (Gekko & Timasheff, 1981a,b). We have shown that the presence of glycerol increased the refolding efficiency of the Cys→Ser 3SS derivatives (Sawano et al., 1992). To study the effect of glycerol on the temperature dependence of refolding reaction, the efficiency of refolding at various glycerol concentration and temperatures was examined for Δ1 (Figure 6). With the increase in glycerol concentration up to 30%, the temperature−activity profile shifted to the higher temperature side with a concomitant increase in the maximal activity at T_{opt}. With a further increase in glycerol concentration, the profile further shifted to the higher temperature side, but the maximal activity did not increase. The optimal

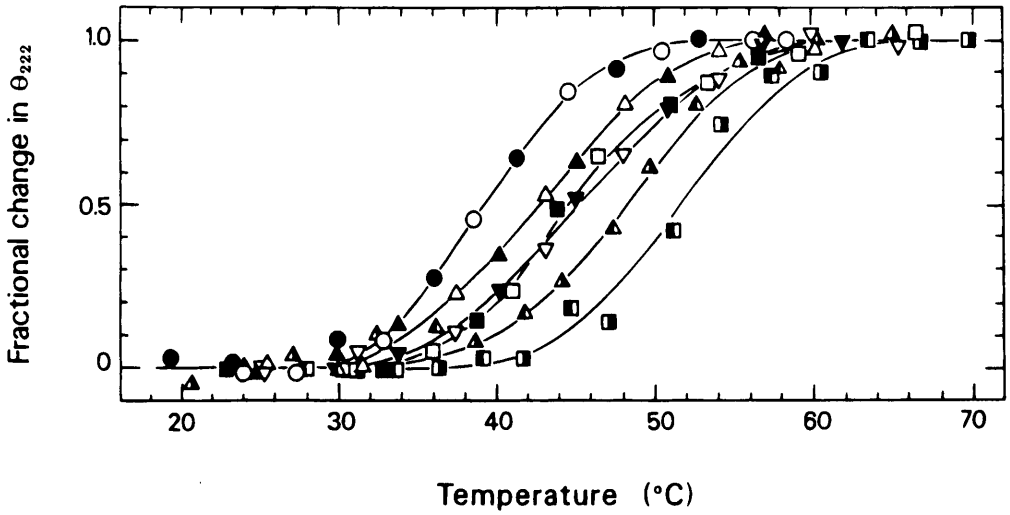


FIGURE 4: Thermal denaturation–renaturation transition curves for the 3SS derivatives. The ordinate shows fractional changes in the mean residue ellipticity at 222 nm. Symbols are as follows (the first ones are for the data obtained on increasing temperature changes, and the second ones are for decreasing temperature changes): (○, ●) Δ1; (△, ▲) Δ2; (▲, ▲) Δ2Ala; (▽, ▼) Δ3; (□, ■) Δ4; (◨, ◧) Δ4Ala.

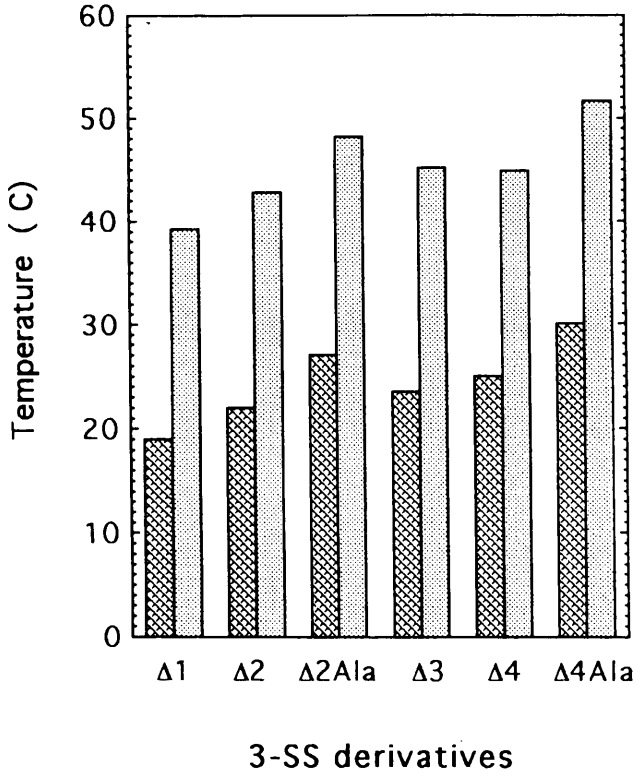


FIGURE 5: Relationship between T_{opt} (cross-hatched) and T_m (dotted) for the 3SS derivatives.

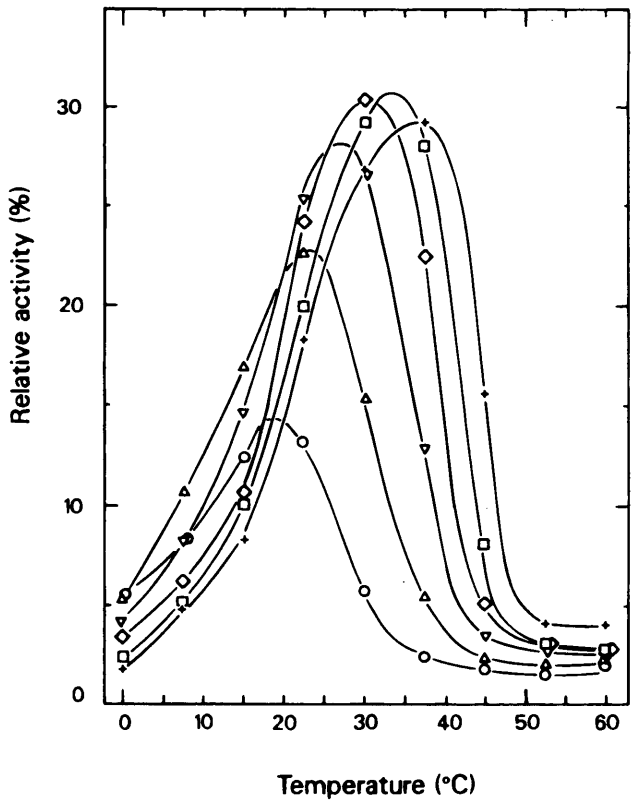


FIGURE 6: Lytic activity of the solution of Δ1 refolded at various temperatures in the absence or presence of glycerol: 0 (○), 10 (△), 20 (▽), 30 (◇), 40 (□), and 50% (+) glycerol (v/v).

temperatures were 19.0, 23.3, 27.0, 30.5, 33.8, and 36.8 °C for the glycerol concentrations of 0, 10, 20, 30, 40, and 50%, respectively. The increment in T_{opt} was 3–4 °C for the increment of 10% in glycerol concentration.

The qualitative features described above can be expressed in a different way if, in Figure 6, we follow the data points for a common temperature value and different glycerol concentrations: at moderately high temperature, say at 37.5 °C, the refolding efficiency was markedly enhanced from nearly zero to 30% by the addition of glycerol while at lower temperature, say at 0 °C, it decreased with the increase in glycerol concentration. Overall, the optimal conditions for the refolding of Δ1 were 30–35 °C in temperature and 30–45% in glycerol concentration.

The HPLC elution profiles of the derivatives reoxidized in the presence of glycerol showed the main peak, whose retention time nearly coincided with that of authentic lysozyme and whose height correlated with the lytic activity of the refolded, crude protein solution (not shown) in a similar way as described before for the case without glycerol.

Then, the refolded and purified Δ1 was subjected to the CD thermal denaturation experiments in the presence of glycerol (Figure 7). Reversible transition curves were obtained in the presence of glycerol, and the transition temperature shifted to the higher side with increasing concentration of glycerol. T_m's were 39.2 (Figure 4), 45.6, and 51.6 ± 0.6 °C for glycerol concentrations of 0, 20, and

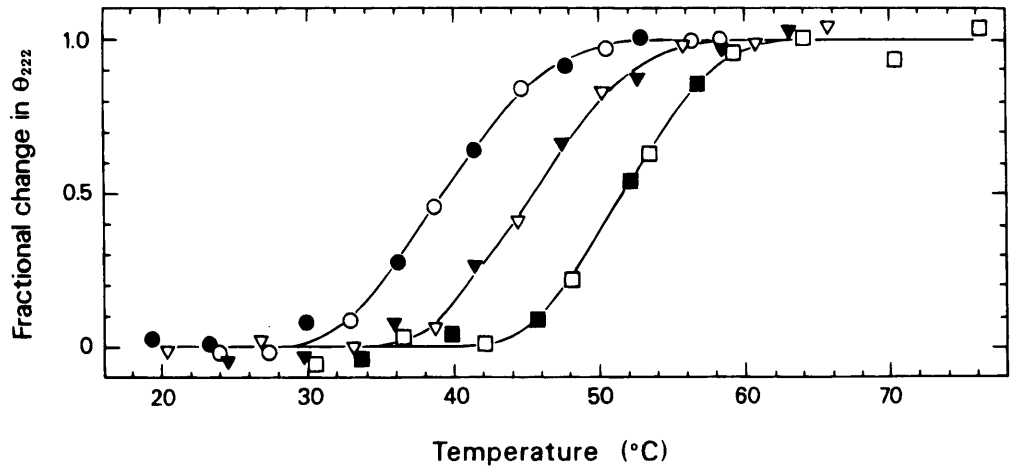


FIGURE 7: Thermal denaturation−renaturation transition curves of Δ1 in the absence (○, ●) or presence of 20% (▽, ▼) or 40% (□, ■) of glycerol. The ordinate shows fractional changes in the mean residue ellipticity at 222 nm. Open and filled symbols indicate the data obtained during increasing and decreasing temperature changes, respectively.

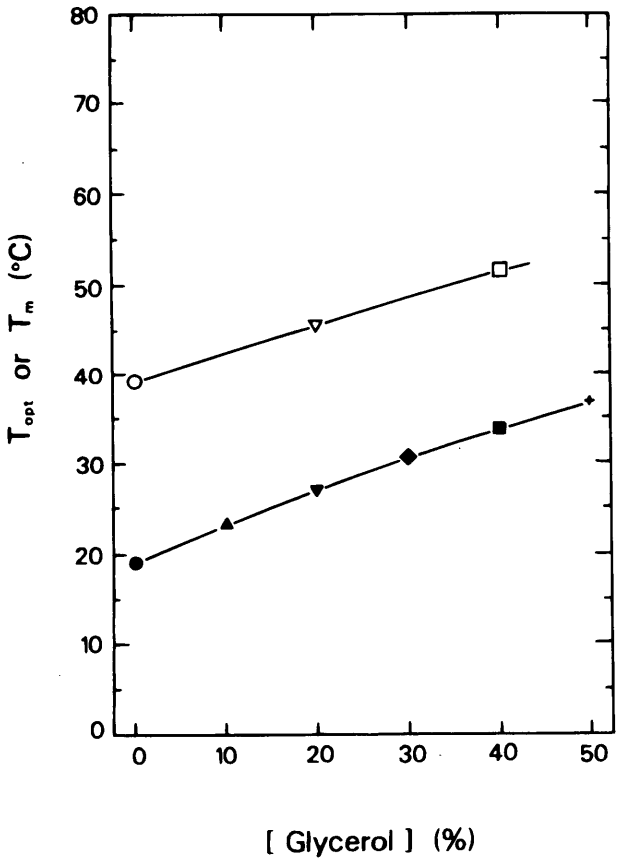


FIGURE 8: Relationship between T_{opt} (lower line) and T_m (upper line) of Δ1 for various concentration of glycerol. The same symbols as in Figures 5 and 6 are used for each glycerol concentration.

40%, respectively. The increment in T_m was 3.2−3.0 °C per 10% increment in glycerol concentration. This value was nearly equal to the increment in T_{opt} described before, and a nearly parallel relationship between T_m and T_{opt} was evident as shown in Figure 8. The differences between T_m and T_{opt} were 20.2, 18.6, and 17.8 °C for glycerol concentrations of 0, 20, and 40%, respectively.

DISCUSSION

Parallelism between T_{opt} and T_m. In this study, it was shown for hen lysozyme 3SS derivatives that there exists a parallel relationship between the optimal temperature for oxidative refolding and the denaturation temperature of the completely refolded state. A set of disulfide bridge-engineered proteins used here was suitable for the demonstration of the relationship with sufficient accuracy: none of the four disulfide bridges is necessary for the refolding of hen lysozyme (Acharya & Taniuchi, 1977; Sawano et al., 1992), and the 3SS derivatives all refolded, although with different efficiencies, into the state that had enough enzymatic activity for the determination of refolding optima as well as the secondary and tertiary structure, which could undergo reversible thermal transition. Such a relationship as demonstrated here will be potentially useful in the optimization of the refolding temperature for expressed, engineered proteins.

The existence of T_{opt} in the "recovered activity"−"refolding temperature" profile and a sharp decrease in the recovered activity with an increase in the refolding temperature are phenomenologically similar to the existence of T_{opt} in the activity−temperature profile of enzymes and probably result from the two factors which have mutually opposing temperature dependence: one is the increase in the rates of elementary reaction steps of oxidative refolding with the increase in temperature, and the other is the destabilization, with the increase in temperature, of the species on the refolding pathway. Here, depending on the stability of the completely refolded state, two cases can be considered: In the first case, the completely refolded species is sufficiently stable, and the reverse reaction through the rate-determining step from the completely refolded species to (an) intermediate species is negligible. In this case, the temperature dependence of the refolding rate will be governed by the stability of the intermediate species. In the second case, the completely oxidized and correctly refolded species is not sufficiently stable, and its reduction, coupled with denaturation, is not negligible. Then, the temperature dependence of the refolding rate as well as the yield will be affected by the stability of the completely oxidized and refolded species.

The oxidative refolding reaction of the hen lysozyme 3SS derivatives studied here was in accord with the second case as described in the Results section. The reaction step leading to the completely refolded Δ1 was shown to be reversible under the present refolding conditions. The rough agreement between the fraction of the recovered activity at 30 °C to

the optimal one (Figure 1) and the fraction of the intact Δ1 remaining after the 50-min incubation at 30 °C (Figure 3) suggests that the major determinant of the temperature dependence of the refolding of Δ1 is the stability of the completely refolded state. The existence of a parallel relationship between T_{opt} and the T_m for the completely refolded states of the 3SS derivatives is therefore considered to be characteristic of the refolded proteins, which have low stability against thermal as well as reduction-coupled denaturation. (Here, we must also note that T_m as well as T_{opt} is generally a function of the concentration of thiol reagents.)

On the other hand, the former case has been clearly demonstrated for the regeneration, i.e., oxidative refolding, of RNase A by Rothwarf and Scheraga (1993a). They showed that RNase A rengerated through multiple pathways and that the rate-determining step involved a rearrangement of 3SS species (Rothwarf & Scheraga, 1993b). They also showed that the regeneration pathways of RNase A depended on the nature of the redox reagent used (Rothwarf & Scheraga, 1993c), and observed correspondingly different temperature dependence of the regeneration rates (Rothwarf & Scheraga, 1993d), which indicated that the temperature dependence is governed by the stability of an intermediate species. It was identified as the 3SS species that was markedly populated at 15 °C. For the oxidative refolding of bovine pancreatic trypsin inhibitor (Creighton, 1977) and that of α-LA (Ewbank & Creighton, 1993a,b), the rate-determining steps have been shown to be an intramolecular two-disulfide rearrangement and that of a two- and three-disulfide rearrangement, respectively.

For authentic (i.e., 4SS) hen lysozyme, it has been shown that the oxidative refolding proceeds through a limited search of intermediates (Ristow & Wetlaufer, 1973; Anderson & Wetlaufer, 1976): two disulfide bonds involving Cys64, Cys76, Cys80, and Cys94 being formed rapidly, followed by the disulfide bond between Cys30 and Cys115, and presumably finally, the disulfide bond between Cys6 and Cys127. This outline of oxidative refolding is nearly the reverse of the reduction pathway for α-LA reported by Ewbank and Creighton (1993a) when the correspondence of the homologous disulfide bridges between the two proteins is taken into consideration, and it may be suggested that the oxidative refolding pathways of the two proteins are roughly similar. However, there also exist differences: In the oxidative refolding of α-LA, multiple one-, two-, and three-disulfide intermediates were generated in contrast to the case for lysozyme; $R^{6/120cam}$, a reduced α-LA in which Cys residues 6 and 120 were blocked with iodoacetamide, did not refold into native (i.e., the remaining three disulfide bridges being in native combinations) state (Ewbank & Creighton, 1993a), whereas in lysozyme, none of the four disulfide bridges is obligatory for the correct folding (Acharya & Taniuchi, 1977; Sawano et al., 1992). At present, unlike in the cases for RNase A, bovine pancreatic trypsin inhibitor, and α-LA due to the lack of the studies for lysozyme that extensively deal in a quantitative way with the kinetic progress of the intermediates in the oxidative refolding pathway, we have little information about either the species involved in a rate-determining step for the oxidative refolding of lysozyme or for that of the refolding of lysozyme 3SS derivatives. It should be noted, however, that the T_m for authentic lysozyme at pH 4 is about 78 °C (Pfeil & Privalov, 1976) while the T_{opt} for its oxidative refolding is 37 °C (Saxena & Wetlaufer, 1970) or 35 °C (this study), resulting in the difference of about 40 °C between T_m and T_{opt}. This value is much larger than the difference of about 20 °C between them for the lysozyme 3SS derivatives studied here. The stability of some intermediate species, most likely 2SS or 3SS, which is much lower than that of the completely refolded authentic lysozyme (=4SS), probably governs the T_{opt}. In other words, the species produced immediately after the rate-determining step will be rapidly stabilized by the formation of the third and/or fourth disulfide bridge(s), making the reduction-coupled denaturation of the completely oxidized and refolded 4SS species almost negligible in the temperature range around, and moderately higher than, T_{opt}.

Stability Differences among the 3SS Derivatives. A simple explanation for the destabilization of protein due to the loss of a disulfide bridge is the increase in the entropy of the denatured state, which is brought about by the change in the organization of covalently closed loops formed by main chain and disulfide bridges. In accordance with this notion, Cooper et al. (1992) have shown that the reduction in T_m resulting from the removal of the 6−127 disulfide bond in hen lysozyme 3SS derivative $CM^{6,127}$ is attributed totally to an increase in the entropy difference between the native and denatured states. In the present case, the entropy increase in the denatured state of each 3SS derivative compared to that of the 4SS form was estimated, based on the equation derived by Lin et al. (1984), to be 17.6, 16.1, 14.5, and 14.8 cal/K·mol for Δ1−Δ4, respectively. If there is no change between 3SS and 4SS forms in the enthalpy as well as the entropy difference for the unfolding reaction except for this entropy increase, the amount of the decrease in T_m with the removal of a single disulfide bridge is predicted to be $\Delta 1 > \Delta 2 > \Delta 4 \gtrsim \Delta 3$. The present result of CD T_m measurement agreed with this prediction. However, the change between the 3SS and the 4SS forms in the enthalpy difference as well as in the entropy difference for the denaturation reaction was observed in our preliminary differential scanning calorimetry experiment (Tachibana, Oka, Fukada, Takahashi, Sorai, unpublished results). Also, Kuroki et al. (1992) have reported that the destabilization of the disulfide-engineered human lysozyme, which corresponds to Δ4 studied here, is not explained by the change in entropy difference but rather by the change in enthalpy difference. Complete interpretation of the stability change among the 3SS derivatives in terms of the structural features of the lysozyme molecule awaits the result of a detailed thermodynamic investigation.

The Cys→Ala 3SS derivatives (Δ2Ala, Δ4Ala) had higher T_m than their counterpart Cys→Ser derivatives (Δ2, Δ4). The environment of the residue positions in authentic lysozyme where Cys residues are replaced in these 3SS derivatives (30, 115, 76, 94) is inside the molecule (Shrake & Rupley, 1973). Therefore, the stabilization of the Ala derivatives relative to the Ser derivatives can be qualitatively explained by the relationship between the stability of the engineered protein and the hydrophobicity of substituted residues (Yutani et al., 1987). The difference in T_m observed in this study, about 6 °C, however is much smaller than that (27 °C) reported for a single disulfide bond derivative of the bovine pancreatic trypsin inhibitor with the remaining Cys residues replaced by either Ser or Ala (Darby et al., 1991; van Mierlo et al., 1991; Staley & Kim, 1992; the value was corrected for the effect of the substitution or the addition of residues

other than Cys→Ser/Ala replacement according to the description in Staley and Kim). We must note that, in the case of human lysozyme 3SS derivatives, a Cys77,95→Ser mutant has been reported to be more stable by 2.3 °C in T_m than a Cys77,95→Ala mutant, probably due to a hydrogen bond between the two Ser residues (Yamada et al., 1994).

Refolding Efficiency and the Effect of Glycerol. The maximal refolding efficiency shown in Figure 1, evaluated by apparent lytic activity of the reoxidized (i.e., not purified) protein solution, was significantly higher for Δ4, Δ2Ala, and Δ4Ala than for Δ1, Δ2, and Δ3. The same situation is observed for the refolding efficiency at almost any refolding temperature. The reason of the high apparent activity for Δ4 is partly explained by the high activity of purified Δ4 (Sawano et al., 1992). Although the activities of the purified Δ1, Δ2, and Δ3 were 60−90% of the authentic lysozyme, the activity of the purified Δ4 was 130−140% of authentic lysozyme. [This high activity may be related with the proximity of the disulfide bridge Cys76−Cys94, which is removed in Δ4, to Trp62. The replacement of Trp62 with Tyr, Phe, or His has been shown to exhibit enhanced bacteriolytic activity (Kumagai & Miura, 1989), and it was suggested that the smaller size of the aromatic ring at the 62nd position may favor the lytic process.] In the case of Δ2Ala, the activity of the purified protein (about 72% of authentic lysozyme) was similar to that of Δ2 (60−76%). Therefore, the apparently higher activity of the reoxidized solution of Δ2Ala must be related with an increased number of correctly refolded protein molecules compared to Δ2. Likewise, the similar level of refolding efficiency between Δ4 and Δ4Ala combined with the lower activity of the purified Δ4Ala compared to that of the purified Δ4 indicates an increased number of correctly refolded Δ4Ala compared to Δ4. Altogether, the Cys→Ala derivatives studied in this study had a higher refolding efficiency as well as a higher stability than their counterpart Cys→Ser derivatives.

The appearance of significant refolding yield at a progressively higher temperature side with increasing glycerol concentration in the reoxidation reaction of Δ1 (Figure 5) can be explained by the stabilization, during reoxidation reaction, of a completely refolded state with glycerol. The observed amount of stabilization by glycerol of the refolded and purified Δ1 (3.2−3.0 °C per 10% increment in glycerol concentration) as well as the increase in T_{opt} with the increase in glycerol concentration (4−3 °C per 10% increment in glycerol concentration) were significantly higher than that reported for chymotrypsinogen (1.1−0.5 °C), RNase (2.4−1.7 °C) (Gekko & Timasheff, 1981b), and authentic lysozyme (0.7−1.5 °C) (Gekko, 1982). We also carried out the control refolding experiment for authentic lysozyme in the presence of glycerol (not shown). The increase in T_{opt} was about 2 °C per 10% increment in glycerol concentration and was significantly lower than that for Δ1. The enhanced amount of stabilization for Δ1 can be explained as follows: It has been shown that the folded state of protein is stabilized by glycerol with the mechanism of preferential hydration (Gekko & Timasheff, 1981a). In the mechanism, the unfavorable interaction between glycerol and protein tends to minimize the surface of contact between them, thereby stabilizing the compact, folded state relative to the extended, unfolded state. In the present case, the compactness of the folded state of Δ1 is considered to be nearly the same as that of authentic lysozyme since the amount of secondary structures, that of enzymatic activity, and the elution position on RPHPLC were nearly the same with those of authentic lysozyme. On the other hand, the unfolded state of Δ1, which has three disulfide bridges, is naturally expected to be less compact compared to that of authentic lysozyme, which has four disulfide bridges. Therefore, the amount of free energy increase for the unfolded state due to the unfavorable interaction with glycerol is expected to be higher for Δ1 than for authentic lysozyme, resulting in an increased amount of stabilization of the folded state of Δ1 compared to that of authentic lysozyme.

ACKNOWLEDGMENT

We thank Profs. K. Nakamae and Y. Nishizuka for encouragement and support, Prof. K. Akasaka for discussions, critical reading of the manuscript, and the facilities for CD measurements, and Mr. T. Oka for the preparation of Δ4.

REFERENCES

Acharya, A. S., & Taniuchi, H. (1976) *J. Biol. Chem. 251*, 6934−6946.
Acharya, A. S., & Taniuchi, H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2362−2366.
Acharya, A. S., & Taniuchi, H. (1982) *Mol. Cell. Biochem. 44*, 129−148.
Anderson, W. L., & Wetlaufer, D. B. (1976) *J. Biol. Chem. 251*, 3147−3153.
Cooper, A., Eyles, S. J., Radford, S. E., & Dobson, C. M. (1992) *J. Mol. Biol. 225*, 939−943.
Creighton, T. E. (1977) *J. Mol. Biol. 113*, 275−293.
Darby, N. J., van Mierlo, C. P. M., & Creighton, T. E. (1991) *FEBS Lett. 279*, 61−64.
Ewbank, J. J., & Creighton, T. E. (1993a) *Biochemistry 32*, 3677−3693.
Ewbank, J. J., & Creighton, T. E. (1993b) *Biochemistry 32*, 3694−3707.
Gekko, K. (1982) *J. Biochem. 91*, 1197−1204.
Gekko, K., & Timasheff, S. N. (1981a) *Biochemistry 20*, 4677−4686.
Gekko, K., & Timasheff, S. N. (1981b) *Biochemistry 20*, 4667−4676.
Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
Heinrikson, R. L., & Meredith, S. C. (1984) *Anal. Biochem. 136*, 65−74.
Hirs, C. H. W. (1967) *Methods Enzymol. 11*, 199−203.
Kumagai, I., & Miura, K.-I. (1989) *J. Biochem. 105*, 946−948.
Kuroki, R., Inaka, K., Taniyama, Y., Kidokoro, S., Matsushima, M., Kikuchi, M., & Yutani, K. (1992) *Biochemistry 31*, 8323−8328.
Lin, S. H., Konishi, Y., Denton, M. E., & Scheraga, H. A. (1984) *Biochemistry 23*, 5504−5512.
Miki, T., Yasukochi, T., Nagatani, H., Furuno, M., Orita, T., Yamada, H., Imoto, T., & Horiuchi, T. (1987) *Protein Eng. 1*, 327−332.
Mitraki, A., & King, J. (1989) *Bio/Technology 7*, 690−697.
Nagai, K., & Thøgersen, C. (1987) *Methods Enzymol. 153*, 461−481.
Pfeil, W., & Privalov, P. L. (1976) *Biophys. Chem. 4*, 23−32.
Radford, S. E., Woolfson, D. N., Martin, S. R., Lowe, G., & Dobson, C. M. (1991) *Biochem. J. 273*, 211−217.
Ristow, S. S., & Wetlaufer, D. B. (1973) *Biochem. Biophys. Res. Commun. 50*, 544−550.
Rothwarf, D. M., & Scheraga, H. A. (1993a) *Biochemistry 32*, 2671−2679.

Rothwarf, D. M., & Scheraga, H. A. (1993b) *Biochemistry 32*, 2680−2789.
Rothwarf, D. M., & Scheraga, H. A. (1993c) *Biochemistry 32*, 2690−2697.
Rothwarf, D. M., & Scheraga, H. A. (1993d) *Biochemistry 32*, 2698−2703.
Sawano, H., Koumoto, Y., Ohta, K., Sasaki, Y., Segawa, S., & Tachibana, H. (1992) *FEBS Lett. 303*, 11−14.
Saxena, V. P., & Wetlaufer, D. B. (1970) *Biochemistry 9*, 5015−5023.
Shrake, A., & Rupley, J. A. (1973) *J. Mol. Biol. 79*, 351−371.
Staley, J. P., & Kim, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1519−1523.
Tachibana, H., Sasaki, Y., Sawano, H. & Kitakawa, M. (1990) *Protein Eng. 3*, 371.
Taniyama, Y., Yamamoto, Y., Nakao, M., Kikuchi, M., & Ikehara, M. (1988) *Biochem. Biophys. Res. Commun. 152*, 962−967.
van Mierlo, C. P. M., Darby, N. J., Neuhaus, D., & Creighton, T. E. (1991) *J. Mol. Biol. 222*, 373−390.
Yamada, H., Kanaya, E., Ueno, Y., Ikehara, M., Nakamura, H., & Kikuchi, M. (1994) *Biol. Pharm. Bull. 17*, 612−616.
Yutani, K., Ogasahara, K., Tsujita, T., & Sugino, Y. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4441−4444.

Article No. jmbi.1999.3442 available online at http://www.idealibrary.com on IDEAL®

J. Mol. Biol. (2000) 00, 1–14

JMB

The Transition State in the Folding-Unfolding Reaction of Four Species of Three-disulfide Variant of Hen Lysozyme: The Role of Each Disulfide Bridge

Atsushi Yokota[1], Kazuki Izutani[1], Masanori Takai[1], Yukiko Kubo[1] Yasuo Noda[1], Yasuko Koumoto[2], Hideki Tachibana[3] and Shin-Ichi Segawa[1]*

[1]*Department of Physics, School of Science, Kwansei Gakuin University, Nishinomiya, 662-8501, Japan*

[2]*Department of Cell Biology National Institute for Basic Biology, Okazaki, 444-8585, Japan*

[3]*Department of Biology Faculty of Science, and Graduate School of Science and Technology, Kobe University Kobe, 657-8501, Japan*

**Corresponding author*

The effects of lacking a specific disulfide bridge on the transition state in folding were examined in order to explore the folding-unfolding mechanism of lysozyme. Four species of three-disulfide variant of hen lysozyme (3SS-lysozyme) were prepared by replacing two Cys residues with Ala or Ser: C6S/C127A, C30A/C115A, C64A/C80A and C76A/C94A. The recombinant hen lysozyme was studied as the standard reference containing four authentic-disulfide bridges and the extra N-terminal Met the recombinant hen lysozyme containing the extra N-terminal. Folding rates were measured by monitoring the change in fluorescence intensity associated with tri-*N*-acetyl-D-glucosamine binding to the active site of refolded lysozyme. It was confirmed that the folding rate of the recombinant hen lysozyme containing the extra N-terminal was the same as that of wild-type lysozyme, and that the folding rate was little affected by the presence of tri-*N*-acetyl-D-glucosamine. The folding rate of C64A/C80A was found to be the fastest and almost the same as that of the recombinant hen lysozyme containing the extra N-terminal, and that of C30A/C115A the second, and that of C6S/C127A the third. The folding rate of C76A/C94A was particularly slow. On the other hand, the unfolding rates which were measured in the presence of tri-*N*-acetyl-D-glucosamine showed the dependence on the concentration of tri-*N*-acetyl-D-glucosamine. The intrinsic unfolding rate in the absence of tri-*N*-acetyl-D-glucosamine was determined by extrapolation. Also in the unfolding rate, C76A/C94A was markedly slower than the others. It was found from the analysis of binding constants of tri-*N*-acetyl-D-glucosamine to C64A/C80A during the unfolding process that the active site of C64A/C80A partly unfolds already prior to the unfolding transition. On the basis of these kinetic data, we suggest that C64A/C80A folding transition can occur with leaving the loop region around SS3 (C64-C80) flexible, while cross-linking by SS4 (C76-C94) is important for the promotion of folding, because it is an indispensable constraint on the way towards the folding transition state.

Keywords: protein folding; lysozyme variants; disulfide bridge; transition state; protein engineering

Abbreviations used: SS1, disulfide bridge between Cys6 and Cys127; SS2, Cys30-Cys115; SS3, Cys64-Cys80; SS4, Cys76-Cys94; 3SS-lysozyme, a three-disulfide variant of hen lysozyme lacking a specific disulfide bridge, in which respective Cys residues are replaced by Ala or Ser residues, C6S/C127A, C30A/C115A, C64A/C80A, C76A/C94A; wtLYZ, wild-type hen lysozyme; metLYZ, recombinant hen lysozyme containing the extra N-terminal Met; rcmLYZ, C6/C127-reduced and carboxymethylated lysozyme; triNAG, tri-N-acetyl-D-glucosamine; GuHCl, guanidine hydrochloride; CD, circular dichroism; HOHAHA, homonuclear Hartmann-Hahn; ITC, isothermal titration calorimetry; DSC, differential scanning calorimetric; ITC, isothermal titration calorimetry.

E-mail address of the corresponding author: shsegawa@kwansei.ac.jp

0022-2836/00/0000001–14 $35.00/0

Introduction

Understanding protein folding-unfolding transition is the key to elucidating how tertiary structures of protein are determined from physical and chemical properties of molecules. Recent investigations suggest that the collapse to a relatively compact globule occurs before the formation of native structure when a fully unfolded polypeptide chain is placed in a solution of very low denaturant concentration (Kuwajima, 1989; Ptitisyn, 1995; Rodder & Colon, 1997). Extensive secondary structure and some tertiary contacts exist within the collapsed state, but the majority of tertiary interactions are not persistent. Experimental studies on α-lactalbumin have suggested that such a molten globule has a native-like secondary structure and a native-like topology in the α-domain, but no significant structure in the β-domain (Baum *et al.*, 1989; Wu *et al.*, 1995). However, some recent studies indicate that the molten globule state is a disordered compact globule (Sali *et al.*, 1993; Creighton, 1997; Qi *et al.*, 1998). In the early folding of lysozyme, it was recently proposed that the transiently accumulating partially folded state with the a-domain already formed is not located on the direct folding pathway, but rather on an alternative slow-side pathway (Wildegger & Kiefhaber, 1997). Thus, the role of the molten globule in protein folding is still controversial. Kinetic studies with NMR spectroscopy that monitor individual residues demonstrated that the close packing characteristic of the native state emerges in a highly cooperative manner following the rapid formation of the molten globule (Balbach *et al.*, 1995). Folding of this type occurs at the late stage of the transition. Although much effort has been expended to characterize the molten globule state, not much attention has been paid to the late stage of folding. We focus our attention on this rate-determining step in the folding transition of lysozyme.

A computer simulation study of island model protein demonstrated that transient intermediate states between the fully folded and unfolded states are inherently unstable and undergo a transition to either of two final states, and that there exists a predominant folding pathway (Segawa & Kawai, 1986). The ensemble of such transient states along the pathway may be represented by the hypothetical transition state. Characterization of the transition state is an essential step in understanding the mechanism of protein folding. The most effective way is based on protein engineering (Fersht, 1995). Previously, the transition state of lysozyme folding-unfolding process was characterized by effects of protein-solvent interactions and substrate analogue binding on folding and unfolding rate constants (Segawa & Sugihara, 1984a,b). This time, effects of the removal of a specific-disulfide bridge were examined. In the absence of such a disulfide bridge, if the two segments connected by the disulfide bridge interact with each other in the transition state, the folding rate would be diminished by the entropic cost of bringing the two ends into proximity. If the two segments were not fixed in the transition state, little influence on the folding rate would be expected; instead the unfolding rate would be significantly increased due to the loss of disulfide constraint in the transition state. Thus, kinetic data can give information about the local structure around the disulfide bridge in the transition state.

We constructed four species of the three-disulfide variant of hen lysozyme (CSS-lysozyme) by replacing two Cys residues with Ser or Ala residues (Sawano *et al.*, 1992; Tachibana *et al.*, 1994), and showed that none of the four native-disulfide bridges is obligatory for the attainment, *in vitro*, of higher-order structures and bacteriolytic activity. Here, circular dichroism (CD) spectra show that the tertiary structure of 3SS-lysozyme is almost the same as that of wild-type lysozyme (wtLYZ). When the folding reaction of 3SS-lysozyme was initiated by rapidly diluting the guanidine hydrochloride (GuHCl) concentration, the far-UV CD and fluorescence signals quickly recovered the intensity characteristic of the native state during the early folding towards the collapsed molten globule, so that the signal associated with the late folding transition was hardly observed. The key to the late folding of lysozyme is the appearance of the active site to bind tri-*N*-acety-D-glucosamine (triNAG). Here, lysozyme folding was followed by monitoring the change in fluorescence intensity associated with triNAG binding to refolded protein.

Folding and unfolding rate constants were determined for all the species of 3SS-lysozyme. Contrary to our expectation, the contribution of SS1(C6-C127), SS2(C30-C115) and SS3 to increasing the folding rate was only a little, while only SS4 significantly increased the folding rate. In general, the 3SS-lysozyme with a slow folding rate had a slow unfolding rate. In particular, C76A/C94A had the slowest folding rate and the slowest unfolding rate, while C64A/C80A the fastest folding and unfolding rates in the presence of 0.24 mM triNAG. C30A/C115A and C6S/C127A had intermediate folding and unfolding rates. On the basis of these kinetic data, the details of the transition state are discussed with respect to the local structure around an individual-disulfide bridge.

Results

Four species of lysozyme variant lacking a specific-disulfide bridge

Hen lysozyme contains four disulfide bridges; SS1 between C6 and C127, SS2 between C30 and C115, SS3 between C64 and C80 and SS4 between C76 and C94. Figure 1 shows the ribbon diagram of hen lysozyme illustrating the positions of four disulfide bridges, SS1 to SS4, which are represented by yellow sticks. The active site is formed

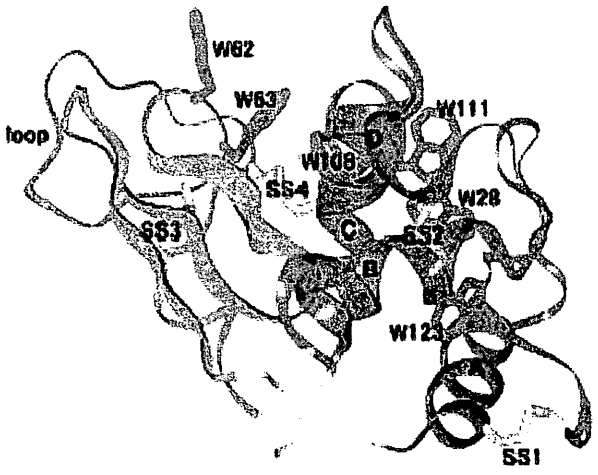


Figure 1. Ribbon diagram of hen lysozyme, produced by Insight II based on the X-ray crystallographic structure (PDB 6LYZ). Four a-helices are colored magenta. Yellow sticks represent four disulfide bridges, SS1 to SS4. SS1 links C-terminal region with A-helix, and SS2 links B-helix with D-helix. The 3_{10}-helix is connected with the triple-stranded b-sheet (colored blue) by SS3. SS4 is located between the loop region and C-helix. The tertiary structure is comprised of two folding domains, one of which is formed from four α-helices (α-domain), and the second from the triple-stranded β-sheet and the 3_{10}-helix (β-domain). The active site is formed at the interface between two domains including three tryptophan residues (colored green, W62, 63 and 108). The other three (W28, 111 and 123) are included in the hydrophobic core region in the α-domain.

at the interface between α and β-domains. Four species 3SS-lysozyme were prepared by means of oxidative-refolding of the reduced form of the variants, each of which lacks a specific pair of cysteinyl residues corresponding to each of the four native-disulfide bridges. Respective Cys residues were replaced by Ala residues, except that Cys6 was replaced by a Ser residue. The variant C6S/C127A lacks SS1. In a similar manner, C30A/C115A, C64A/C80A and C76A/C94A lack SS2, SS3 and SS4, respectively. The peptide maps of these variants were consistent with the attainment of native-disulfide pairings in each of them. Since these 3SS-lysozymes expressed in *Escherichia coli* contain an extra N-terminal Met residue, the recombinant hen lysozyme was prepared and studied as the standard reference containing four authentic disulfide bridges and the extra Met (metLYZ). In addition, another type of 3SS-lysozyme was prepared by the chemical modification of wtLYZ, in which C6 and C127 were replaced by carboxy-methylated Cys residues (rcmLY2). This variant, rcmLYZ, is different from C6S/C127A in the respect of the absence of the N-terminal Met. The amino acid compositions of these lysozyme variants were confirmed by ion-spray mass spectrometry. Obtained masses for C30A/C115A and rcmLYZ were 14374.4 and 14422.9. These values were exactly the same as those calculated from their respective amino acid sequences.

Characterization of the tertiary structures of 3SS-lysozyme

The far-UV CD spectra of metLYZ, rcmLYZ and four species of 3SS-variants were almost equal to those of wtLYZ in the magnitude of major band ellipticities with slight differences in detailed spectral shapes (data not shown). Figure 2 shows near-UV CD spectra for these 3SS-lysozymes measured at pH 4.0 and 4.0 °C. The spectra of wtLYZ and rcmLYZ were almost the same as those of metLYZ shown in Figure 2. For 3SS-variants, the shape of spectra is basically similar to that of metLYZ, but they vary in the magnitude of molecular ellipticity [θ]; for example, the spectrum of C30A/C115A or C76A/C94A is significantly smaller than that of metLYZ in the magnitude of [θ] near 290 nm. Since W28, W111 and W123 surround SS2 (C30-C115), the removal of SS2 seems to alter the environment of these tryptophan residues, especially in respect of local structural flexibility. In a similar manner, W108 is located near SS4 (C76-C94), therefore the near-UV CD spectrum may be reduced remarkably in C76A/C94A. These CD spectra show that each of 3SS-lysozyme has a rigid folded structure besides the native-like secondary structure. Indeed, the near-UV CD spectra disappeared with the unfolding induced by the addition of GuHCl as described later. NMR spectroscopy is more effective to get detailed information about the structure of protein. Homonuclear Hartmann-Hohn (HOHA-HA) spectra were measured for all the species of 3SS-variants. Although the details of NMR data will be published elsewhere, they suggest that all the species of 3SS-lysozyme retain the rigid tertiary structure quite similar to the native one on the

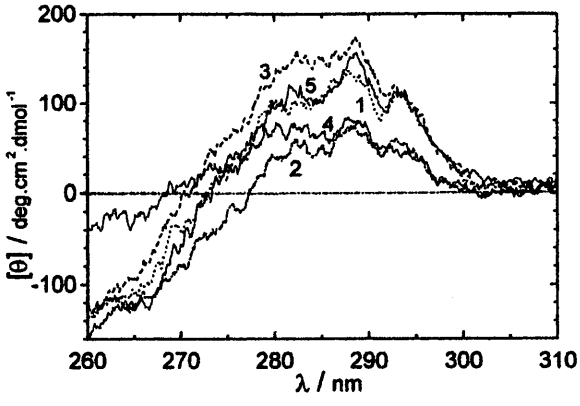


Figure 2. Near-UV CD spectra of 3SS-lysozymes. Spectra were measured in 50 mM sodium acetate buffer (pH 4.0) and 4.0 °C. 1, (···) C6S/C127A; 2, (-·-) C30A/C115A; 3, (---) C64A/C80A; 4, (—) C76A/C94A; 5, (—) metLYZ.

whole, while the local structure around the site of mutation becomes flexible in terms of the protection factor against the amide hydrogen exchange reaction.

The thermal unfolding of these 3SS-lysozymes had been studied by monitoring the change in CD signal at 222 nm (Tachibana *et al.*, 1994). Further, we have carried out the differential scanning calorimetric (DSC) experiments. The van't Hoff enthalpy of unfolding, ΔH_{vH}, is obtained from the DSC curve; $\Delta H_{vH} = 4RT_d^2\Delta C_{ex}/Q$, where T_d is the transition temperature, ΔC_{ex} is the height of the heat absorption peak at T_d, Q is the area of the peak. The ratio of calorimetric enthalpy to the van't Hoff enthalpy ($\gamma = \Delta H_{cal}/\Delta H_{vH}$) were determined (Table 1). They clearly indicate the two-state transition for the thermal unfolding of 3SS-lysozymes. The thermal stability of metLYZ is considerably decreased compared with that of wtLYZ: the decrease in T_d of about 8 deg. C. The difference between two species of lysozyme is only the presence of the N-terminal Met in metLYZ. For the same reason, C6S/C127A is also less stable than rcmLYZ by about 11 °C in T_d. Among all the 3SS-variants, C76A/C94A is the most stable and C6S/C127A the most unstable.

Calorimetric titration of triNAG into 3SS-lysozyme

In order to explore whether the detailed structure of the active site of lysozyme is restored or not in 3SS-variants, binding of a substrate analogue (triNAG) to lysozyme variants was studied using isothermal titration calorimetry (ITC). Titration experiments were performed in 0.1 M sodium acetate buffer (pH 4.0). The titration data were analyzed based on a simple 1:1 binding reaction between lysozyme and triNAG. The stoichiometry (n) was introduced as a curve fitting parameter. Three adjustable parameters, ΔH_b, K_b^0 and n, were determined by non-linear least-squares fitting using an iterative Marquardt algorithm, where K_b^0 and ΔH_b represent the binding constant and enthalpy. Thermodynamic parameters obtained are listed in Table 2. The fact that n is nearly equal to 1.0 is really an indication of good purification of 3SS-variants. The binding constant to C6S/C127A is the largest, while that to C76A/C94A is significantly smaller than the others. Binding enthalpies of 3SS-lysozymes except for C76A/C94A are substantially the same as those of wtLYZ. Although the binding ability of C76A/C94A is somewhat weak in both K_b^0 and ΔH_b, all the species of 3SS-lysozyme can strongly bind the substrate analogue. This indicates that the active site structure of 3SS-lysozyme is quite similar to that of wtLYZ. It should be noted that the lack of a disulfide bridge except for SS4 has little effect on K_b^0 and ΔH_b, nevertheless the tertiary structure itself is greatly destabilized. In other words, as far as lysozyme variants are completely folded, the active site structure is quite similar to that of wtLYZ, but the thermal stability vary among them. The binding ability of C76A/C94A is the weakest but its thermal stability is the highest among all the 3SS-variants. Since SS4 is located at the interface between α and β-domains of lysozyme, the active site of C76A/C94A is likely to deform a little.

Unfolding transition curves of four species of 3SS-lysozyme

The unfolding transition of lysozyme induced by GuHCl was observed monitoring the far-UV CD signal. Figure 3 shows the change in molecular ellipticity at 222 nm ($[\theta]_{222}$) measured at pH 3.0 and 4.0 °C. Transition curves for wtLYZ and metLYZ are shown in Figure 3(a). They were analyzed on the basis of the two-state transition. We assumed the linear relationship between the unfolding free energy [$\Delta G_{U\text{-}N}$] and the concentration of GuHCl [c]:

$$\Delta G_{U\text{-}N} = -RT \ln K_{uf} = m(c_{1/2} - c) \qquad (1)$$

where $c_{1/2}$ is the concentration of GuHCl at the midpoint of the transition curve, K_{uf} the equilibrium constant of unfolding. In the analysis of these transition curves, the fitting parameters were $c_{1/2}$, m and the baselines of $[\theta]_{222}$ in the folded and unfolded states, where m is an indicator of the sharpness of unfolding transition (cooperativity). Curve fitting was carried out on the assumption that the baseline of $[\theta]_{222}$ is a linear functions of c with two fitting parameters (the slope and intercept). In the folded state, the slope of the baseline was fixed zero. The best fitting values for m and $c_{1/2}$ are listed in Table 3. The transition curve of rcmLYZ monitored by the change in $[\theta]_{222}$ is

Table 1. Transition temperatures and the ratio of $\Delta H_{cal}/\Delta H_{vH}$ for lysozyme unfolding

	T_d (°C)	pH	γ	T_d (°C)	pH	γ
wtLYZ	68.2	3.1	0.98	75.9	3.7	1.02
metLYZ	59.5	3.0	1.01	68.7	4.0	1.01
rcmLYZ	39.6	3.0	0.98	48.8	4.0	0.98
C6S/C127A	28.1	3.0	1.03	36.4	3.5	1.00
C30A/C115A	38.5	3.0	0.99	47.9	3.9	1.01
C64A/C80A	37.5	3.0	1.01	47.2	3.9	0.99
C76A/C94A	41.5	2.9	0.95	51.3	3.9	0.97

$\gamma = \Delta H_{cal}/\Delta H_{vH}$.

Table 2. triNAG binding to 3SS-variants of lysozyme

	T (°C)	$K_b^0 \times 10^{-5}$ (M^{-1})	n	ΔH_b (kJ/mol)
wtLYZ	9.7	1.5	1.00	−45
	19.7	1.3	1.06	−49
metLYZ	19.7	0.94	1.01	−44
rcmLYZ	9.1	2.4	1.08	−45
C6S/C127A	9.1	3.0	0.96	−44
C30A/C115A	9.5	1.3	0.98	−50
C64A/C80A	9.1	1.6	0.99	−45
C76A/C94A	9.0	0.64	0.97	−26

ITC experiments were carried out at pH 4.1.

shown in Figure 3b. It is noted that the baseline $[q]_{222}$ of the unfolded state of rcmLYZ markedly slants upward from 2 M to 4 M of GuHCl concentrations. This characteristic is common to all the species of 3SS-lysozyme as shown in Figure 3(c) and (d). Probably, this means that there remain secondary structures to a considerable extent in the unfolded state of these 3SS-lysozymes, because the concentration of GuHCl is relatively low. In order to confirm that the tertiary structure of rcmLYZ fully unfolds at 2.0 M GuHCl, the near-UV CD spectra were observed at various GuHCl concentrations. Figure 3(b) shows the change in $[\theta]_{288}$ measured at pH 3.0 and 4.0 °C. The unfolding transition of the tertiary structure of rcmLYZ is indeed observed near 1.0 M GuHCl. The analysis of this transition curve was carried out in a similar manner. Obtained best fitting parameters were

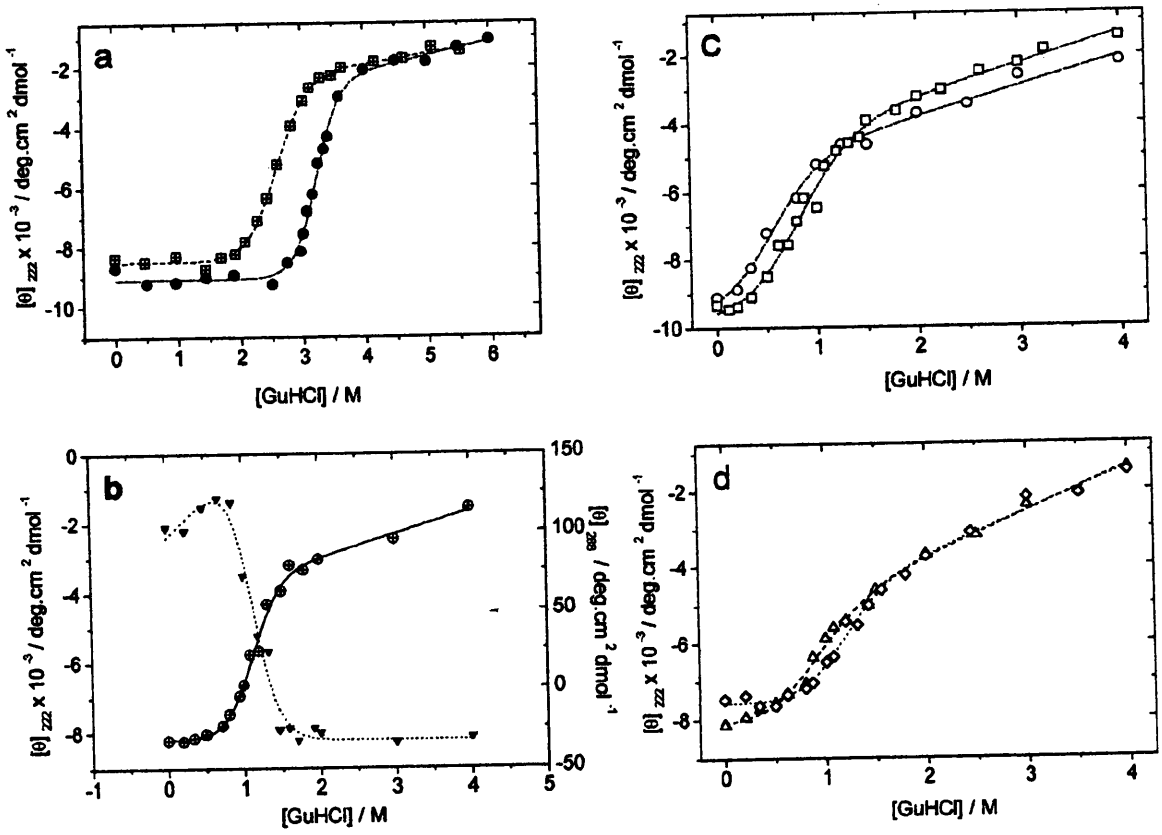


Figure 3. Unfolding transition curves of lysozyme induced by GuHCl. The change in $[\theta]_{222}$ or $[\theta]_{288}$ was monitored. (a) Unfolding transition curves of wtLYZ, (●); metLYZ (□). (b) The unfolding transition curve of rcmLYZ based on the change in $[\theta]_{222}$, (○); $[\theta]_{288}$ (c) Unfolding transition curves of C6S/C127A, (○); C30A/C115A, (□). (d) Unfolding transition curves of C64A/C80A, (△); C76A/C94A, (◇). The baselines of $[\theta]_{222}$ were assumed as a linear function of GuHCl concentration in both folded and unfolded states. The best-fitting curves obtained by non-linear least-squares fitting are shown in each panel.

Table 3. Unfolding transition of lysozyme derivatives induced by GuHCl

	$c_{1/2}$ (M)	m (kJ mol^{-1}M^{-1})	$m\,c_{1/2}$ (kJ/mol)
wtLYZ	3.22	13	42
metLYZ	2.58	9.7	25
rcmLYZ	1.10	13	14
C6S/C127A	0.51	9.2	4.7
C30A/C115A	0.77	8.3	6.4
C64A/C80A	0.83	8.7	7.2
C76A/C94A	1.12	11	12

The best fitting baselines are shown in Figure 3.

$m = 14$ kJ mol^{-1}M^{-1} and $c_{1/2} = 1.09$ M, which were quite close to those obtained from the transition curve of $[\theta]_{222}$.

Figure 3(c) and (d) show transition curves for 3SS-lysozymes. In these cases, the analysis of the transition curve was somewhat difficult because C6S/C127A or C30A/C115A unfolds though a slight at 0 M GuHCl. The baseline was assumed to be flat in the folded state. The best fitting parameters, m and $c_{1/2}$, are listed in Table 3. According to the unfolding-free energy at 0 M GuHCl ($m\ c_{1/2}$), C76A/C94A is the most stable and C6S/C127A the most unstable among all the species of 3SS-lysozyme. This is consistent with the result for their thermal stability.

The concentration-jump of GuHCl was carried out in order to observe kinetics of refolding and unfolding of lysozyme variants. In most cases, folding processes may be followed monitoring the change in CD or fluorescence signal. However, the signal amplitude associated with folding generally diminishes rapidly with decreasing GuHCl concentrations under the folding condition. Particularly, the signal amplitude associated with folding of 3SS-variants was extremely small, because folding experiments had to be performed at GuHCl concentrations below 0.8 M. In order to follow the folding process, therefore, we observed the change in fluorescence intensity due to triNAG binding to refolded lysozyme. The fluorescence intensity of tryptophan residues in lysozyme is quenched by the protonation of certain carboxyl groups (probably E35, D101 and D52) in the complex with triNAG, that is to say, the complex formation causes the loss of fluorescence in the acidic pH region (Lehrer & Fasman, 1967). Indeed, a marked change in fluorescence intensity could be observed associated with refolding of 3SS-variants in the presence of triNAG. But it was necessary for us to examine the effect of triNAG binding on refolding kinetics of lysozyme. In the absence of triNAG, fortunately the change in fluorescence intensity was still observable in the folding processes of wtLYZ, metLYZ, rcmLYZ and C76A/C94A, although it was quite small. For these species, therefore, folding rates were able to be determined both in the absence and presence of triNAG. As described below, both rate constants were confirmed to coincide with each other.

Folding of wtLYZ, metLYZ and rcmLYZ

Figure 4 shows the time course of fluorescence intensity during the refolding process of metLYZ (pH 3.0 and 4.0 °C) caused by the concentration-jump from 4.0 M to 1.3 M GuHCl, in the absence and presence of triNAG. Line (a) represents fluorescence intensity in an initial unfolded state at 4.0 M GuHCl. Curve (b) shows the refolding process in the absence of triNAG, and Curve (c) in the presence of 0.24 mM triNAG. The signal amplitude associated with folding in the absence of triNAG was very small under the folding condition. The presence of triNAG greatly increased the signal amplitude, but the time constant was almost the same. When a drop of concentrated triNAG solution was added to a solution containing metLYZ fully refolded at 1.3 M GuHCl, the fluorescence intensity rapidly decreased due to triNAG binding to metLYZ, but no detectable relaxation process

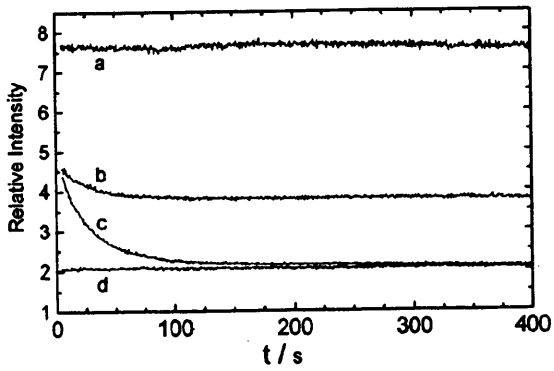


Figure 4. The time course of fluorescence intensity of metLYZ upon the GuHCl concentration jump from 4.0 M to 1.3 M. Refolding condition was pH 3.0 and 4.0 °C. The excitation wavelength was 282 nm (a bandwidth of 1.5 nm) and the emission wavelength 370 nm (a bandwidth of 20 nm). The protein concentration was 1.7 μM in all experiments shown in this figure. (a) An initial fluorescence intensity of unfolded metLYZ. (b) Refolding process observed in the absence of triNAG. The signal amplitude was very small (the time constant about 29 seconds). (c) Refolding process in the presence of 0.24 mM triNAG (the time constant 31 seconds). (d) triNAG binding to metLYZ refolded beforehand at 1.3 M GuHCl.

was observed as shown by the line (d). This indicates that the binding process is rapid enough to finish in the dead time of mixing. Therefore, triNAG binding observed under the refolding condition certainly reflects the time course of increasing population of fully folded lysozyme. The time constant of folding was little affected by the presence of triNAG; it was 29 seconds in the absence of triNAG and 31 seconds in the presence of 0.24 mM triNAG in Figure 4. In addition, the agreement between two folding rates was confirmed at other concentrations of GuHCl, and also in cases of wtLYZ and rcmLYZ, as shown in Figure 5(a). This result suggests that building the active site of lysozyme is later than the rate-determining step in lysozyme folding. This is consistent with our previous conclusion for wtLYZ (Segawa & Sugihara, 1984b). Further, the change in CD signal at 222 nm or 288 nm was monitored associated with refolding in the absence of triNAG. As shown in Figure 5(a), the folding rates agree well with those determined from the time course of fluorescence, although the ratio of signal to noise was much worse in the CD signal change.

The folding rates of metLYZ were identical with those of wtLYZ, as shown in Figure 5(a). The presence of N-terminal Met has little effect on the folding rate. On the other hand, the folding rate of rcmLYZ was found to be slower by a factor of 1/5 than that of wtLYZ or metLYZ at concentrations of GuHCl around 1.0 M. However, two other groups have reported that the former is rather twice faster at pH 6.7, 0.5 M GuHCl (Denton *et al.*, 1994) or comparable to the latter at pH 4.5 and 0.27 M GuHCl (Eyles *et al.*, 1994). These results seem to contradict our result, but these refolding experiments were performed under strongly native conditions. The final concentrations of GuHCl were varied in our experiments. The folding rates of rcmLYZ come close to those of metLYZ, as GuHCl concentrations decrease. Therefore, our result does not necessarily conflict with previous ones.

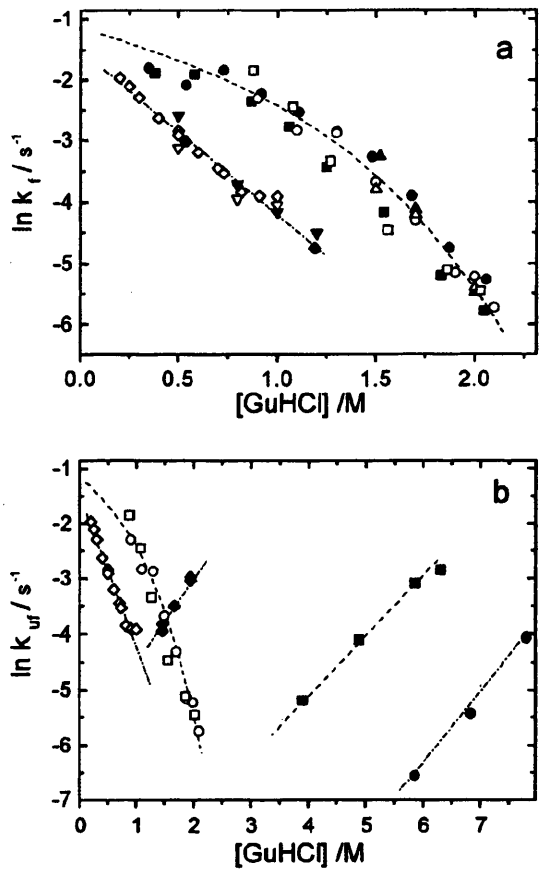


Figure 5. (a) Folding rates of wtLYZ, metLYZ and rcmLYZ. Rate constants were determined from the change in fluorescence intensity or CD signal. wtLYZ fluorescence, (○, ●); $[\wedge]_{222}$, (▲); $[\theta]_{288}$, (△). metLYZ, fluorescence, (□, ■). rcmLYZ: fluorescence, (◇, ◆), $[\theta]_{222}$, (▲), $[\theta]_{288}$ (△). In cases of fluorescence, open symbols represent the folding rates in the absence of triNAG, and filled symbols those in the presence of 0.24 mM triNAG. (b) Unfolding rates of wtLYZ, metLYZ and rcmLYZ. Filled symbols represent the unfolding rates: wtLYZ, (●); metLYZ, (■); rcmLYZ, (◆). Open symbols represent the folding rates shown in the upper panel.

Folding of 3SS-variants of lysozyme

Folding processes of 3SS-variants of lysozyme were followed monitoring the change in fluorescence after the concentration jump from 2.0 M GuHCl to a desired final concentration. For example, Figure 6 shows the refolding process of C64A/C80A at 0.24 M GuHCl, pH 3.0 and 4.0 °C. In the absence of triNAG, a rapid decrease in fluorescence intensity was observed within a dead time of mixing, after that only a slight change was observed associated with protein folding as shown by curve (b) of Figure 6. Curve (c) represents the refolding process in the presence of 0.24 mM triNAG, and curve (d) triNAG binding to C64A/C80A fully refolded beforehand at 0.24 M GuHCl. This shows that triNAG binding to folded C64A/C80A is much faster than C64A/C80A folding. The small decay in curve (d) is likely to reflect the slight increase in population of folded protein induced by the addition of triNAG. It was difficult to determine precisely the time constant of curve (b). In the cases of 3SS-variants, therefore, it was indispensable to perform refolding experiments in the presence of triNAG. In the same manner, refolding rates were determined for all the 3SS-variants at various GuHCl concentrations. In order to confirm that the folding rate is independent of the concentration of triNAG, folding rates were determined in the presence of various concentrations of triNAG. For example, the folding rates of C64A/C80A were also determined in the presence of 10 mM triNAG, and found to be almost the same

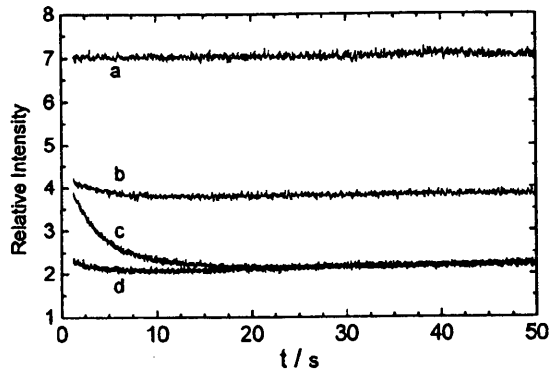


Figure 6. The time-course of fluorescence intensity of C64A/C80A upon the GuHCl concentration jump from 2.0 M to 0.24 M. Refolding condition was pH 3.0 and 4.0 °C. The excitation wavelength was 282 nm (a bandwidth of 1.5 nm) and the emission wavelength 350 nm (a bandwidth of 20 nm). The protein concentration was 1.4 μM in all experiments shown in this figure. (a) An initial fluorescence intensity of unfolded C64A/C80A. (b) Refolding process observed in the absence of triNAG. (c) Refolding process in the presence of 0.24 mM triNAG (the time constant 3.5 seconds). (d) triNAG binding to C64A/C80A refolded beforehand at 0.24 M GuHCl.

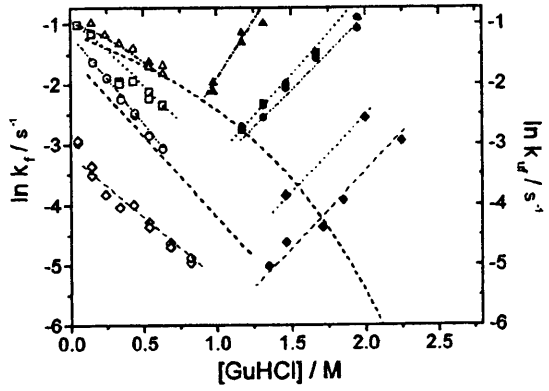


Figure 7. Folding and unfolding rates of 3SS-variants of lysozyme which were determined in the presence of 0.24 mM triNAG. Open symbols represent the folding rates of each 3SS-lysozyme: C6S/C127A, (○); C30A/C115A, (□); C64A/C80A, (△); C76A/C94A, (◇). Two broken lines represent the best-fitting curves for folding rates of metLYZ and rcmLYZ shown in Figure 5. Unfolding rates of each 3SS-lysozyme were determined in the presence of 0.24 mM tirNAG and are shown by filled symbols (the same as those of the folding rates). Unfolding rates of C76A/C94A could be determined in the absence of triNAG, and are shown in this figure (◇).

as those in the presence of 0.24 mM triNAG. The results are shown in Figure 7. The folding rate of C64A/C80A is the fastest among all the species of 3SS-lysozyme and nearly equal to that of metLYZ. Second to C64A/C80A, the folding rate of C30A/C115A is the fastest, and C6S/C127A is after C30A/C115A. The folding rate of rcmLYZ is nearly equal to that of C6S/C127A. The N-terminal Met residue of C6S/C127A has little influence on the folding rate of 3SS-lysozyme, as well as in folding kinetics of metLYZ. The folding rate of C76A/C94A is remarkably slowed down. The variants C64A/C80A and C76A/C94A are in a marked contrast, in spite of the fact that SS3 and SS4 are located close to each other.

Unfolding of 3SS-lysozyme

Unfolding rates for wtLYZ, metLYZ and rcmLYZ were determined in the absence of triNAG. They were measured by monitoring the change in fluorescence intensity. The results are shown in Figure 5(b). The unfolding rate of metLYZ is faster than that of wtLYZ by a factor of about 30. The N-terminal Met residue significantly increases the unfolding rate. For the same reason, C6S/C127A unfolds faster than rcmLYZ by a factor of about 50 (Figure 7). Probably, in the folded state, the antiparallel β-sheet between residues 1-3 and 38-40 may be strongly affected by the extra Met residue. On the other hand, 3SS-variants of lysozyme unfold at GuHCl concentrations around 1.0 M, and unfolding rates were too fast for kinetic measurements by manual mixing. Therefore, unfolding rates of 3SS-lysozymes except for C76A/C94A could not be observed in the absence of triNAG; the time constants were less than 2.0 seconds. Since the unfolding rate of C76A/C94A was slow enough to follow kinetics by manual mixing, it was observed both in the absence and presence of triNAG (Figure 7). In the presence of triNAG, the unfolding reaction generally proceeds through two pathways: unfolding of free lysozyme dissociated from triNAG (k_0) and unfolding of lysozyme associated with triNAG (k_1). The apparent unfolding rate constant, k_a, is represented by:

$$k_a = k_0/(\alpha + 1) + \alpha k_1/(\alpha + 1), \quad \alpha = K_b^G L_0 \quad (2)$$

where K_b^G is the triNAG binding constant to lysozyme remaining transiently in the folded state during the unfolding process, and L_0 the analytical concentration of triNAG. Apparent rate constants were determined from unfolding experiments in the presence of various concentrations of triNAG. Unfolding rates observed in the presence of 0.24 mM triNAG are plotted in Figure 7 together with folding rates at the same concentration of triNAG. These data clearly indicate that C64A/C80A is the fastest in the unfolding rate, C6S/C127A or C30A/C115A the second and C76A/C94A the slowest. Unfolding rates of C76A/C94A were markedly slower than the others.

Since k_a is strongly affected by L_0, the intrinsic unfolding rate (k_0) should be compared with each other to consider the unfolding process of 3SS-lysozyme. For the determination of k_0 and K_b^G, the reciprocals of apparent unfolding rates ($\tau_a = 1/k_a$) were plotted against L_0:

$$\tau_a = [1/k_0 + (K_b^G/k_0)L_0]/[1 + (k_1 K_b^G/k_0)L_0] \quad (3)$$

When $(k_1 K_b^G/k_0)\, L_0 \ll 1$, the relationship between τ_a and L_0 is linear, as shown in Figure 8. This condition is satisfied in most cases of low concentrations of L_0, because k_1 is much less than k_0. Therefore, K_b^G/k_0 is easily determined from the slope of the straight line in Figure 8, and $1/k_0$ from the intercept at $L_0 = 0$. The unfolding rates of C76A/C94A and rcmLYZ could be determined also in the absence of triNAG. In these cases, the good linear relationship between τ_a and L_0 was confirmed in the vicinity of $L_0 = 0$. Non-linear least-squares fitting for τ_a was required for the determination of another parameter k_1. The values of best fitting parameters are listed in Table 4.

From kinetic data at 1.46 M GuHCl, k_0 of C30A/C115A is the fastest. Next to it, k_0 of C64A/C80A and C6S/C127A are similar to each other. Also in the intrinsic unfolding rate, C76A/C94A was much slower than the others.

Here we mention the values of K_b^G. They are smaller than K_b^0 listed in Table 2 by a factor of about 1/10. This is not surprising because K_b^G is the binding constant measured at a high concentration of GuHCl (1.46 M) during unfolding. But it should be noted that the ratio of K_b^G/K_b^0 for C64A/C80A is extremely small compared with the others. Probably, this means that the active site of C64A/C80A partly unfolds already in the folded state during the unfolding process. This may be a kind of intermediate observed on the way of C64A/C80A unfolding. The loop region around SS3 is rich in hydrophilic residues. In the absence of SS3, it is likely that guanidinium ions permeate this region to destroy the local structure prior to the unfolding transition.

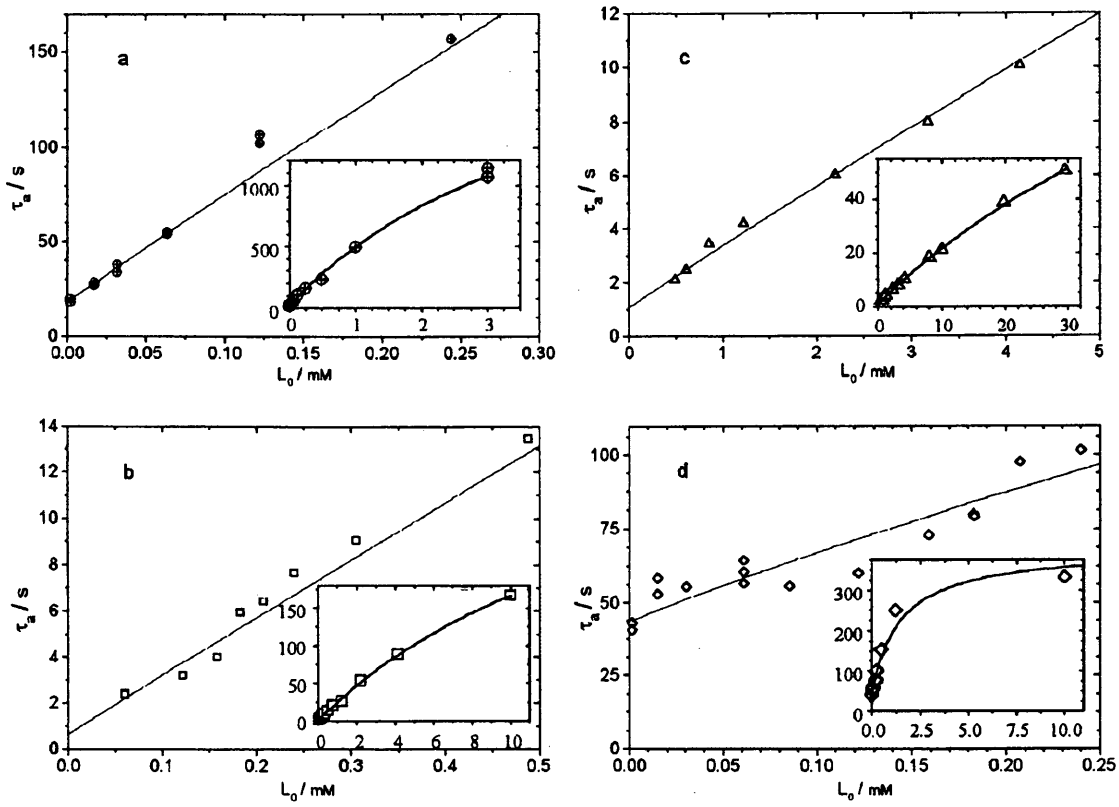


Figure 8. The dependence of τ_a on triNAG concentrations. (a) rcmLYZ: (b) C30A/C115A, c: C64A/C80A: and (d) C76A/C94A. A lot of data of τ_a was obtained at lower concentrations of L_0. The linear relationship exits between τ_a and L_0. The inset of each figure shows the dependence of τ_a on higher concentrations of L_0. The best-fitting curves were obtained by non-linear least-squares fitting with parameters k_0, k_1 and K_b^G. In the cases of rcmLYZ and C76A/C94A, τ_a were determined in the absence of triNAG and are plotted in (a) and (d). The best fitting parameters are listed in Table 4.

Table 4. Unfolding process of 3SS-lysozymes

	$K_b^G \times 10^{-4}$ (M^{-1})	K_b^G/K_b^0	k_0 (s^{-1})	k_1 (s^{-1})
C6S/C127A	2.9	0.10	0.99	0.004
C30A/C115A	3.8	0.29	1.5	0.002
C64A/C80A	0.21	0.01	0.91	0.005
C76A/C94A	0.63	0.10	0.022	0.002
rcmLYZ[a]	3.2	0.13	0.054	0.0004

Unfolding condition: pH 3.0, 4.0 °C, 1.46 M GuHCl.
[a] 1.95 M GuHCl only in rcmLYZ unfolding.

Discussion

The heterogeneity in the compact disordered state

In the case of wtLYZ folding (four disulfide bridges intact), the formation of a partially folded intermediate state was observed within about 30 ms following an initial hydrophobic collapse. This state was investigated in detail using the pulsed hydrogen exchange method (Radford *et al.*, 1992). The experiments revealed that the a-domain of lysozyme is partly folded in this state, while the β-domain remains unstructured. However, this intermediate state was not detected in rcmLYZ refolding (Eyles *et al.*, 1994; Denton *et al.*, 1994). These data suggest that the early intermediate with non-native tertiary interactions is destabilized in the refolding process of rcmLYZ. In our kinetic experiments on 3SS-lysozyme folding, therefore, the early intermediate state appears to be destabilized in the compact disordered state following the burst change in fluorescence. However, in general, the partially folded intermediate state and fully disordered state may coexist through kinetic partitioning in the compact disordered state. Indeed, in the early stage of wtLYZ folding, the triangular folding mechanism was proposed among the intermediate (I), unfolded (U) and native (N) states (Wildegger & Kiefhaber, 1997). According to this model, there are two folding pathways, and it was found that the partially folded intermediate state is not located on the direct folding pathway, but rather on an alternative slow side pathway. Further this analysis pointed out that the absolute free energy of the transition state in the rate-determining step is identical on both folding pathways. These results imply that there are some alternative pathways to acquire the final tertiary structure, but the nature of the rate-determining step leading to the final state is unaltered; different pathways might merge into a single transition state just near the native state. It was proposed that the transition state of lysozyme folding is quite close to the native one from the fact that the change in heat capacity associated with the unfolding process is nearly equal to zero (Segawa & Sugihara, 1984a). Our experiments on the late stage of 3SS-lysozyme folding were discussed on the hypothesis that the folding of 3SS-lysozymes occurs going through a single transition state from the compact disordered state.

The ϕ-value in the transition state

Figure 9 shows the free energy diagram of lysozyme folding. N-state, U-state and T-state represent the native, unfolded and transition states of lysozyme, respectively. I-state denotes the metastable compact disordered state rapidly formed during the refolding process. $\Delta G_{N\text{-}U}$ represents the difference in free energy between N and U-states; $\Delta G_{N\text{-}U} = G_N - G_U$. In a similar manner, $\Delta G_{T\text{-}U}$ and $\Delta G_{I\text{-}U}$ are the difference in free energy between two states. In order to examine the effect of the removal of a disulfide bridge on the reaction diagram, the ϕ-value is introduced (Matouschek *et al.*, 1989):

$$\phi_T = \delta\Delta G_{T\text{-}U}/\delta\Delta G_{N\text{-}U} = 1 - (RT/\delta\Delta G_{N\text{-}U})\ln(k'_{uf}/k_{uf}) \quad (4)$$

$$\phi_I = \delta\Delta G_{I\text{-}U}/\delta\Delta G_{N\text{-}U} = \phi_T + (RT/\delta\Delta G_{N\text{-}U})\ln(k'_f/k_f) \quad (5)$$

Where $\delta\Delta G_{T\text{-}U}$ is the difference in $\Delta G_{T\text{-}U}$ between metLYZ and a 3SS-variant: $\delta\Delta G_{T\text{-}U} = \Delta G'_{T\text{-}}$

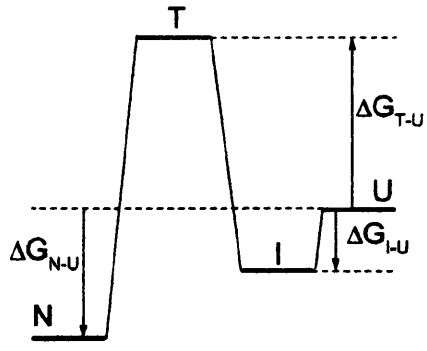


Figure 9. Free energy diagram of 3SS-lysozyme folding. T-state denotes the transition state, I-state the compact disordered state, N-state the native state and U-state the unfolded state. ΔG represents the difference in Gibbs free energy between two states.

$_U - \Delta G_{T\text{-}U}$. Hereafter, δ means the difference between wild-type and mutant proteins, and Δ means the difference between two states, and a prime indicates a quantity of 3SS-variants. $\delta\Delta G_{N\text{-}U}$ is also given by the change in free energy of each state caused by the mutation:

$$\delta\Delta G_{N\text{-}U} = \delta G_N - \delta G_U$$

where

$$\delta G_N = G'_N - G_N, \quad \delta G_U = G'_U - G_U$$

In the unfolded state, the removal of a disulfide bridge greatly increase the conformational entropy of polypeptide chain, accompanying the change in solvation energy of the amino acid side-chain. On the other hand, in the folded state, the overall structure of 3SS-variants is preserved, but the local structure around the site of mutation is likely to become more flexible. Such an increase in entropy tends to be compensated by the loss of internal interaction energy. It is likely that δG_U is a large negative value, and that δG_N is a small positive or negative value. The ϕ-value is always 1.0 in the folded state and zero in the unfolded state. If $\phi_T = 0$, then $\delta G_T = \delta G_U$. This indicates that the conformational multiplicity of the local structure around the site of mutation significantly increases in the transition state as well as in the unfolded state. On the other hand, if $\phi_T = 1.0$, then $\delta G_T = \delta G_N$. In this case, the effect of mutation on the transition state is quite similar to that on the native state. This indicates that the transition state is located near the native state between the native and the unfolded states.

In order to determine the ϕ-value, both k_{uf} and k_f must be known. Unfortunately, in our experiments, both rate constants could not be determined at the same GuHCl concentration. Unfolding rates for 3SS-variants were obtained only at 1.46 M GuHCl. Therefore, only the ϕ_T-value was determined at 1.46 M GuHCl. Under this unfolding condition, I-state is not really observed because U-state is more stable than metastable I-state. The unfolding rate constant of metLYZ at 1.46 M GuHCl was estimated from Figure 5(b), and the value for $\Delta G_{N\text{-}U}$ was given by equation (1). According to equation (4), the ϕ_T-values were obtained for all the species of 3SS-variants. The results are shown in Table 5. The ϕ-value in the transition state is nearly equal to zero for three variants: C6S/C127A, C30A/C115A and C64A/C80A, while the ϕ_T for C76A/C94A is markedly large. The polypeptide conformation is expected to become close to the native one in the transition state, paying a large amount of entropy cost. Nevertheless, the magnitude of δG_T is nearly equal to δG_U in the case of the removal of SS1, SS2 or SS3. This indicates that the removal of these disulfide bridges significantly increases the multiplicity of the local conformation around the site of mutation in the transition state as well as in the unfolded state. Only in the case of the removal of SS4, ϕ_T is about 0.37. It is not straightforward to interpret the intermediate ϕ-value. When $\phi_T = 0.37$, $\delta G_T = 0.37\delta G_N + 0.63\delta G_U$. Apart from the term of $0.37\delta G_N$, roughly speaking, the conformational multiplicity released by the removal of SS4 is limited to 63 % of that released in the unfolded state. But this estimation is too rough, because the value of δG_N is unknown. The ϕ_T value of 0.37 indicates that the effect of the removal of SS4 on the transition state is close to that on the native state though in part, and that the local structure around the site of mutation is not relaxed so much as in the unfolded state.

Table 5. ϕ-values for 3SS-lysozymes

	ϕ_T	ϕ_T^*
C6S/C127A	0.087	0.078
C30A/C115A	−0.13	0.027
C64A/C80A	−0.08	0.0
C76A/C94A	0.37	0.43

ϕ_T-values were obtained at 1.46 M GuHCl.
$\phi_T^* = \phi_T - \phi_I$, ϕ_T^*-values were obtained at 0.25 M GuHCl.

Under the strongly folding condition, I-state should be considered, because the fast folding process towards the compact disordered state exists. However, unfolding rate constants could not be determined under this condition. Since folding rate constants could be observed for all the species of 3SS-variants and metLYZ at 0.25 M GuHCl, the difference of ϕ_T from ϕ_I can be obtained according to equation (5). Although the ϕ-value is unknown in the compact disordered state, we tried to obtain the value of $\phi_T^* = \phi_T - \phi_I$ in order to examine the position of the transition state between I-state and N-state. The results are shown in Table 5. Also under this folding condition, ϕ_T^*-value is nearly equal to zero for 3SS-variants except for C76A/C94A. Only in the case of C76A/C94A, the ϕ_T^*-value is markedly large. If $\phi_T^* = 0$, then $\delta G_T = \delta G_I$. This means that the effect of mutation on the transition state is quite similar to that on the compact disordered state. Under both folding and unfolding conditions, consequently, the removal of SS1, SS2 or SS3 significantly increases the multiplicity of the local conformation around the site of mutation in the transition state as well as in the unfolded or compact disordered state. Only in the case of the removal of SS4, the effect of mutation on the transition state is close to that on the native state though in part.

The role of each disulfide bridge

As mentioned above, the ϕ-value in the transition state of 3SS-variants except for C76A/C94A is nearly equal to zero. This indicates that the local disordered conformations around the site of mutation still persist in the transition state as well as in the unfolded state. That is to say, the folding transition can occur with leaving the region flexible, and the ordered structure around the site of

mutation (SS1, SS2 or SS3) may be recovered after passing through the transition state. On the other hand, in the folding process of C76A/C94A lacking SS4, the conformational multiplicity released by the removal of SS4 is lost in part during the process towards the transition state. Since SS4 is located at the interface between α and β-domains unlike SS1, SS2 (α-domain) and SS3 (β-domain), the interface region may be required to close for the folding transition even in the absence of SS4.

Let us consider the role of each disulfide bridge in an authentic lysozyme. Cross-linking by SS3 introduced to metLYZ is favorable for folding due to the reduction of the conformational multiplicity in the compact disordered state, but rather unnecessary for folding to restrict the rearrangement of polypeptide chain in the transition state. This is the reason why the contribution of SS3 cross-linking towards the promotion of folding is small. The contribution of SS2 or SS1 to increasing the folding rate is also a little. Particularly at lower concentrations of GuHCl, the folding rates of 3SS-variants lacking SS1 or SS2 are quite similar to those of metLYZ. The role of SS1 and SS2 in the folding process is rather close to that of SS3. On the other hand, only SS4 significantly contributes towards the promotion of folding. It may be indispensable for the folding transition to close the interface region around SS4 even in the absence of SS4. This means that SS4 cross-linking is an indispensable constraint on the way towards the transition state, unlike SS3 cross-linking is an unnecessary constraint in the transition state.

Unfolding rates are greatly increased by the removal of a specific disulfide bridge. But a significant difference is found in the unfolding rate among four species of 3SS-variants. In particular, C76A/C94A has the slowest unfolding rate, while C64A/C80A the fastest unfolding rate in the presence of 0.24 mM triNAG. The barrier of free energy is greatly diminished for C64A/C80A unfolding. This is entirely consistent with the above-mentioned interpretation that the folding transition of C64A/C80A occurs with leaving the local structure around SS3 disordered. In other words, the role of SS3 introduced to metLYZ is to prevent lysozyme from unfolding due to restrictions on the transition state. The internal interactions among amino acid residues around SS3 may be insufficient to preserve the local structure in the absence of SS3, therefore the covalent constraints on the transition state are required. On the other hand, in the case of C76A/C94A folding, it was suggested that the interface region around SS4 between α and β-domains seems to close in the transition state, even in the absence of SS4. The multiplicity of the local conformation around the site of mutation may be restricted in the transition state of C76A/C94A unfolding. This is the reason why the increase in unfolding rate by the removal of SS4 is relatively small. The unfolding transition of C76A/C94A must occur with leaving the interface region between two domains around SS4 relatively ordered.

Effects of the slow proline cis-trans isomerization

Here we need to mention the slow phase of refolding process. Folding kinetics of protein including proline residues are generally complicated by the slow interconversion of proline *cis-trans* isomers in the unfolded state. In the case of lysozyme, the time constant of the slow phase is supposed to be nearly equal to 70 seconds, and its signal amplitude is approximately 10% of the total one (Kato *et al.*, 1982). Actually, it is quite difficult to distinguish the slow phase from the major refolding process whose signal amplitude is 10 times larger, as shown in Figure 4. Fortunately, however, folding kinetics of lysozyme is little affected by the presence of the slow phase of refolding, because the fraction of incorrect proline isomers is approximately 10% (the details were discussed by Segawa & Sigihura, 1984a). Therefore, we analyzed folding kinetics as a single exponential function, and did not try to distinguish the slow phase from the major process.

Materials and Methods

Hen egg white lysozyme recrystallized six times, triNAG was purchased from Seikagaku Kogyo Co. Ltd. IPTG, glutathione oxidized and reduced form (GSSG, GSH), reduced DTT, GuHCl were purchased from Wako Pure Chemicals Industries Ltd.

Three-disulfide variants of hen lysozyme

The expression and purification of three-disulfide variants of hen lysozyme have been published (Tachibana *et al.*, 1994). Four species of 3SS-lysozyme used in the present experiments were C6S/C127A, C30A/C115A, C64A/C80A and C76A/C94A, among which for C6S/C127A and C64A/C80A the variant genes were constructed anew in this study in a similar way as described for C30A/C115A and C76A/C94A. Peptide map and N-terminal analyses for the purified, reduced form of these variants showed that they all retained the N-terminal Met residue unexcised in *E. coli*. The recombinant hen lysozyme (metLYZ) was prepared, which is the same as the wild-type hen lysozyme except for containing an extra N-terminal Met. Reoxidized C6S/C127A, C30A/C115A, C64A/C80A and metLYZ were purified by reversed-phase HPLC on a μ-bondashpere C4 column (19 mm × 150 mm, Waters) with a linear gradient of acetonitrile from 28% to 40% in 0.05% (v/v) TFA at the flow rate of 3.0 ml/minute. Purified protein in the main peak fraction eluted as a single peak when rechromatographed on the same column for analytical use. The purification of C76A/C94A was carried out on a TSK TMS-250 column (4.6 mm × 75 mm, Tosoh). The eluted 3SS-lysozyme was concentrated on Sep-pak plus CN cartridge column (Waters), and freeze-dried. Concentrations of the 3SS-lysozymes were estimated by using $A_{280} = 2.64$ per mg/ml protein.

C6/C127-reduced and carboxymethylated lysozyme (rcmLYZ) was prepared essentially according to the

method of Acharya & Taniuchi (1980). Hen egg white lysozyme (3 mg/ml in 0.125 M Tris-acetate buffer (pH 8.1)) was incubated with 5 mM DTT at 23 °C. Reduction was quenched by the addition of fivefold molar excess of iodoacetic acid. The soluble fraction was collected by centrifugation, then applied to the gel filtration column (Sephadex G25, Pharmacia) and eluted in the 0.1 N acetic acid, and freeze-dried. Chemically modified lysozyme was further purified by ion exchange chromatography (Hitrap SP prepacked column, Pharmacia) in 0.05 M sodium phosphate buffer (pH 7.0). rcmLYZ and unmodified lysozyme were separated with a linear gradient of NaCl from 0 Mto 1.0 M. The eluted rcmLYZ was dialyzed at 4 °C against water adjusted to pH 3.8 with HCl and freeze-dried.

Mass spectrometry

The ion-spray mass spectra of C30A/C115A and rcmLYZ were measured on a PE Sciex API 300 triple quadrupole mass spectrometer. Protein in 50% (v/v) aqueous acetonitrile containing 0.05% (v/v) formic acid was infused at 2 µl/minute through the ion-spray interface. The obtained masses for C30A/C115A and rcmLYZ were 14374.4 and 14422.9. These value were exactly the same as those calculated from their respective amino acid sequences.

Isothermal titration calorimetry

Calorimetric experiments were performed with an OMEGA titration calorimeter from MicroCal, Inc., Northampton, MA. Lysozyme variants dissolved in 0.1 M sodium acetate buffer (pH 4.0) was contained in the reaction cell and stirred at 400 rpm. A 250-µl injection syringe was filled with triNAG dissolved in the same buffer. Typically, initial concentrations of lysozyme variants and triNAG were 30 µM and 1.5 mM. 4.0 µl of triNAG solution was regularly injected into reaction cell. The heat effects associated with dilution of triNAG solution into buffer were negligibly small.

CD and fluorescence spectroscopy

CD spectra of 3SS-lysozyme were obtained on a Jasco J-600 spectropolarimeter (Japan Spectroscopic Co. Ltd.) with a thermostatically controled cell holder. Fluorescence spectra were measured with a Shimadzu RF5300PC fluorescence spectrophotometer. The cell was defrosted with the flow of dry nitrogen gas below the room temperature. Protein concentration for near-UV CD spectra was 30-50 µM, that for far-UV CD 2.5-5.0 µM, and that for fluorescence 1.5-3.0 µM.

Kinetics of folding and unfolding

To monitor folding kinetics of 3SS-lysozyme, the fluorescence and CD measurements were carried out using spectrometers mentioned above. Samples were initially dissolved to a protein concentration of 60-80 µM (fluorescence measurements) or 25-45 µM (CD measurements) in 50 mM glycine-HCl buffer (pH 3.0) containing 2.0 M GuHCl. Refolding was initiated by a manual dilution into the final refolding buffer contained in the cell thermostated at 4.0 °C. The temperature of the initial sample solution was carefully controlled beforehand in another thermostat not to change the temperature of refolding buffer upon the dilution. The dilution into the final buffer was 1:40 for fluorescence measurements and 1:10 for CD measurements. The dead times of refolding kinetics were about 2.0 seconds (fluorescence measurements) and 4.0 seconds (CD measurements). The final refolding condition was 50 mM glycine (pH 3.0), and the final GuHCl concentrations from 0.05 M to 0.70 M. In the refolding experiments in the presence of triNAG, the desired concentrations of triNAG were contained in a final buffer solution. The folding rates were quite independent of the concentration of triNAG. For the refolding of wtLYZ and metLYZ, the initial unfolding buffer was 4.0 M in GuHCl. The excitation wavelength for fluorescence measurements was 282 nm with a bandwidth of 1.5 nm. The emission wavelength to monitor refolding kinetics was 370 nm (wtLYZ, metLYZ, rcmLYZ), 360 nm (C6S/C127A), 320 nm (C30A/C115A), 350 nm (C64A/C80A) and 360 nm (C76A/C94A) with a bandwidth of 20 nm. To monitor refolding kinetics of wtLYZ, metLYZ, rcmLYZ and C76A/C94A in the absence of triNAG, the ellipticity at 222 nm or at 288 nm was measured with a bandwidth of 2 nm.

Unfolding kinetics were also monitored by the change in fluorescence and CD signals in the similar manner. The initial condition of sample solution was 50 mM glycine buffer (pH 3.0), 0 M GuHCl. Small amount of the solution was quickly diluted into the final unfolding buffer contained in the observation cell; thus the final unfolding condition was the desired GuHCl concentrations (1.0-2.0 M for 3SS-lysozyme, 4.0-6.0 M for metLYZ, 6.0 -8.0 M for wtLYZ) in 50 mM glycine buffer (pH 3.0) and 4.0 °C. In unfolding experiments in the presence of triNAG, the initial sample solution typically contained 10 mM triNAG.

Acknowledgments

We thank Mr H. Sawano for the construction of C6S/C127A and C64A/C80A genes. We are indebted to Biotechnology Research Laboratories, Takara Shuzo Co., Ltd, Otsu, Shiga for the use of the ion-spray mass spectrometer, and thank Masaru Miyagi for performing the mass spectrometry.

References

Acharya, A. S. & Taniuchi, H. (1980). Preparation of a two-disulfide bonded enzymically active derivative from hen egg lysozyme. *Int. J. Peptide Protein Res.* **15**, 503-509.

Balbach, J., Forge, V., van Nuland, N. A. J., Winder, S. L., Hore, P. J. & Dobson, C. M. (1995). Following protein folding in real time using NMR spectroscopy. *Nature Struct. Biol.* **2**, 865-870.

Baum, J., Dobson, C. M., Evans, P. A. & Hanley, C. (1989). Characterization of a partly-folded protein by NMR methods: studies of the molten globule state of guinea pig α-lactalbumin. *Biochemistry*, **28**, 7-13.

Bax, A. (1989). Two-dimensional NMR and protein structure. *Ann. Rev. Biochem.* **58**, 223-256.

Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III & Skelton, N. J. (1996). *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego, CA.

Creighton, T. E. (1997). How important is the molten globule for correct protein folding? *Trends Biochem. Sci.* **22**, 6-10.

Denton, M. E., Rothwarf, D. M. & Sheraga, H. A. (1994). Kinetics of folding of guanidine-denatured hen egg white lysozyme and carboxymethyl (Cys6, Cys127)-lysozyme: a stopped-flow absorbance and fluorescence study. *Biochemistry*, **33**, 11225-11236.

Eyles, S. J., Radford, S. E., Robinson, C. V. & Dobson, C. M. (1994). Kinetic consequences of the removal of a disufide bridge on the folding of hen lysozyme. *Biochemistry*, **33**, 13038-13048.

Fersht, A. R. (1995). Characterizing transition states in protein folding: an essential step in the puzzle. *Curr. Opin. Struct. Biol.* **5**, 79-84.

Kato, S., Shimamoto, N. & Uchiyama, H. (1982). Identification and characterization of the direct folding process of hen egg-white lysozyme. *Biochemistry*, **21**, 38-43.

Kuwajima, K. (1989). The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins: Struct. Funct. Genet.* **6**, 8-103.

Lehrer, S. S. & Fasman, G. D. (1967). Fluorescence of lysozyme and lysozyme substrate complexes. *J. Biol. Chem.* **242**, 4644-4651.

Matouschek, A., Kellis, J. T., Jr, Serrano, L. & Fersht, A. R. (1989). Mapping the transition state and pathway of protein folding by protein engineering. *Nature*, **340**, 122-126.

Ptitsyn, O. B. (1995). Molten globule and protein folding. *Advan. Protein Chem.* **47**, 83-229.

Qi, P. X., Sosnick, T. R. & Englander, S. W. (1998). The burst phase in ribonuclease A folding and solvent dependence of the unfolded state. *Nature Struct. Biol.* **5**, 882-884.

Radford, S. E., Dobson, C. M. & Evans, P. A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature*, **358**, 302-307.

Roder, H. & Colon, W. (1997). Kinetic role of early intermediates in protein folding. *Curr. Opin. Struct. Biol.* **7**, 15-28.

Sali, A., Shakhnovich, E. & Karplus, M. (1994). How does a protein fold? *Nature*, **369**, 248-251.

Sawano, H., Koumoto, Y., Ohta, K., Sasaki, Y., Segawa, S. & Tachibana, H. (1992). Efficient *in vitro*. folding of the three-disulfide derivatives of hen lysozyme in the presence of glycerol. *FEBS Letters*, **303**, 11-14.

Segawa, S. & Kawai, T. (1986). Computer simulation of the folding-unfolding transition of island-model proteins-folding pathways, transition process and fluctuations. *Biopolymers*, **25**, 1815-1835.

Segawa, S. & Sugihara, M. (1984a). Characterization of the transition state of lysozyme unfolding. I. Effect of protein-solvent interactions on the transition state. *Biopolymers*, **23**, 2473-2488.

Segawa, S. & Sugihara, M. (1984b). Characterization of the transition state of lysozyme unfolding II. Effects of the intrachain crosslinking and the inhibitor binding on the transition state. *Biopolymers*, **23**, 2489-2498.

Tachibana, H., Ohta, K., Sawano, H., Koumoto, Y. & Segawa, S. (1994). Relationship between the optimal temperature for oxidative refolding and the thermal stability of refolded state of hen lysozyme three-disulfide derivatives. *Biochemistry*, **33**, 15008-15016.

Wildegger, G. & Kiefhaber, T. (1997). Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate. *J. Mol. Biol.* **270**, 294-304.

Wu, L. C., Peng, Z. Y. & Kim, P. S. (1995). Bipartite structure of the a-lactalbumin molten globule. *Nature Struct. Biol.* **2**, 281-286.

Edited by C. R. Matthews

(Received 16 August 1999; received in revised form 2 December 1999; accepted 2 December 1999)