STUDIES OF THE *Escherichia Coli* CHAPERONIN PROTEIN GroEL (cpn 60)

Graeme James Thomson

Thesis submitted to the University of Glasgow for the degree of Doctor of Philosophy

Department of Biochemistry

May, 1994

For my parents

Contents **Page No.**

Contents i

List of Figures ix

List of Tables xii

Abbreviations xiii

Acknowledgements xv

Summary xvi

Chapter 1

Introduction

1.1 Molecular chaperones: general introduction 1

1.2 Protein folding *in vivo* versus protein folding *in vitro* 2

1.3 The protein folding problem 6

1.4 Catalysis of protein folding 8

1.4.1 Protein disulphide isomerase 8

1.4.2 Peptidyl prolyl *cis-trans* isomerase 10

1.5 Molecular chaperone proteins 11

1.5.1 The Nucleoplasmins 12

1.5.2 The Hsp70 family 13

1.5.2.1 General background 13

1.5.2.2 Conserved structure of the Hsp70 class 14
1.5.2.3 General mechanism for Hsp70 chaperone mediated action 16
1.5.2.4 Regulation of Hsp70 function 17
1.5.3 The Hsp90 family 19
1.5.4 The sHSP family 20
1.5.5 Co-translational chaperones 21
1.5.6 The Hsp60 family:- the chaperonins 22
1.5.6.1 General history 22
1.5.6.2 The *gro* E gene products 23
1.5.6.3 GroE homologues in higher organisms 25
1.5.6.4 Studies on GroE mediated (re)folding 26
1.5.6.5 GroEL as a catalyst 29
1.5.6.6 Recognition of target polypeptides by GroEL 30
1.5.6.7 The stoichiometry of binding and the site of recognition 32
1.5.6.8 GroEL-GroES interactions 33
1.5.6.9 The ATP binding site of GroEL 36
1.6 Sequential action of the chaperone Hsp70 and Hsp60 proteins 41
1.7 Aims of the project 43

Chapter 2

Materials and Methods

2.1 Materials and reagents 44
2.1.1 Chemicals and biochemicals 44
2.1.2 Enzymes and proteins 45
2.1.3 Chromatography media 46
2.1.4 Pre-packed media 46

	2.1.5 Bacterial strains	46
2.2	General laboratory methods	48
	2.2.1 pH measurement	48
	2.2.2 Protein estimation	48
	2.2.3 Lyophilization	48
2.3	SDS-polyacrylamide gel electrophoresis	48
	2.3.1 Staining for protein	49
2.4	Enzyme assays	49
	2.4.1(a) GroEL ATPase coupled continuous assay	49
	2.4.1(b) GroEL ATPase "coupled quench" assay	50
2.5	Growth of *E. coli* DH1/pGT3270 and *E. coli* DH1/pND5	50
	2.5.1 Growth media	50
	2.5.2 Selection supplements	51
	2.5.3 Cell growth	51
	2.5.4 Cell breakage	51
2.6	GroEL purification	52
	2.6.1 Storage of purified GroEL	52
2.7	Determination of GroEL concentration	52
2.8	Unfolding and refolding of GroEL	52
	2.8.1 Concentration of Guanidinium chloride	52
	2.8.2 Unfolding of GroEL	53
	2.8.3 Attempted refolding of GroEL	53
2.9	Characterisation of GdnHCl treated GroEL	53
	2.9.1 CD analysis	53
	2.9.2 Fluorescence studies	54
	2.9.3 Light scattering measurements	54
	2.9.4 GroEL ATPase activity	54
	2.9.5 ANS fluorescence	54
	2.9.6 Proteinase susceptibility	55

2.10 Molecular weight determinations 55

2.11 Chemical modification and inactivation of GroEL ATPase activity 56

2.11.1 4-(iodoacetamido)salicylicacid treatment 56

2.11.2 5'-p-fluorosulfonylbenzoyl adenosine treatment 56

2.11.3 Determination of oATP concentration 56

2.11.4 oATP inactivation reaction 56

2.11.5 Protection by substrate against oATP inactivation 57

2.11.6 Borohydride reduction of oATP treated GroEL 57

2.11.7 Stoichiometry of incorporation of oATP 57

2.12 Limited proteolysis of GroEL 58

2.13 Peptide isolation and sequencing 58

2.13.1 Acid-washed glassware 58

2.13.2 Preparation of oATP modified GroEL 59

2.13.3 Digestion of oATP modified GroEL 59

2.13.4 Reverse phase HPLC of peptides 59

2.13.5 Modified peptide sequence analysis 60

2.13.6 Mass spectroscopy of the modified peptide 60

2.14 Characterisation of proteinase K-treated GroEL 61

2.14.1 Limited proteolysis with proteinase K 61

2.14.2 Electroblotting of proteinase K-treated GroEL 61

2.14.3 Staining and destaining of PVDF membranes 62

2.14.4 N-terminal sequence analysis 62

2.14.5 Stoichiometry of oATP incorporation into proteinase K-treated GroEL 62

2.15 Titration of thiol groups with DTNB 63

Chapter 3

Purification of GroEL

3.1	Introduction	64
3.2	GroEL purification	64
	Step 1: Extraction and centrifugation	64
	Step 2: 'Negative' chromatography on DEAE-cellulose	65
	Step 3: Salt fractionation with $(NH_4)_2SO_4$	65
	Step 4: Sephacryl S-200 gel permeation chromatography	65
	Step 5: Chromatography on Mono Q	65
	Step 6: Chromatography on Alkyl-Superose	66
	Step 7: Superose 6 chromatography	66
3.3	Purity of GroEL	73
3.4	Molecular weight determination of GroEL	73
3.5	Purification yield of GroEL	73
3.6	Discussion	73

Chapter 4

Affinity labelling the GroEL ATPase active site

4.1	Introduction	75
4.2	Affinity labelling with oATP	75
	4.2.1 Background	75
	4.2.2 Criteria for a potential affinity label	76
	4.2.3 Kinetics of oATP inactivation of GroEL ATPase activity	78
	4.2.4 Substrate protection against oATP inactivation	82

4.2.5 Stoichiometry of incorporation of oATP into GroEL 85
4.2.6 Stoichiometry of incorporation of oATP into substrate protected GroEL 85
4.3 Identification of the oATP modified peptide 87
4.4 Peptide mapping 87
4.4.1 Limited proteolysis of GroEL 87
4.4.2 Proteolysis of oATP modified GroEL 88
4.4.3 Differential peptide mapping of oATP modified-GroEL 88
4.4.4 Isolation and purification of the modified peptide 89
4.4.5 Amino acid sequencing of the modified material 97
4.5 Mass spectroscopy analysis of the modified material 97
4.6 Sulphydryl titres of GroEL with 5, 5'-dithiobis(2-nitrobenzoate) 99
4.7 Studies on proteinase K-treated GroEL 104
4.7.1 Stoichiometry of oATP incorporation into proteinase K-treated GroEL 104
4.7.2 Stabilisation effects of oATP against proteolysis with proteinase K 107
4.7.3 Effect of oATP against proteolysis with proteinase K 107
4.8 Potential active site reagents 111
4.8.1 5'-Fluorosulphonylbenzoyl adenosine 111
4.8.2 4-(iodoacetamido)salicylic acid 111
4.8.3 Pyridoxal 5'-phosphate 112
4.8.4 Methyl acetimidate 112
4.9 Discussion 113

Chapter 5

The unfolding and attempted refolding of GroEL

5.1	Introduction	123
5.2	Unfolding of GroFL by GdnHCl	123
	5.2.1 Changes in secondary structure	123
	5.2.2 Changes in tertiary structure	124
	5.2.3 Changes in quaternary structure	132
	5.2.4 Changes in GroEL ATPase activity	132
	5.2.5 Comparison of the various techniques	135
5.3	Refolding of GdnHCl denatured GroEL	137
	5.3.1 Refolding of GroEL after denaturation in low concentrations of GdnHCl	137
	5.3.2 Refolding of GroEL after denaturation in higher concentrations of GdnHCl	137
	5.3.3 Secondary structure changes	138
	5.3.4 Tertiary structure changes	138
	5.3.5 Quaternary structure changes	138
5.4	Characterisation of the 'intermediate' state of GroEL	144
	5.4.1 8-anilino-1-naphthalenesulphonate (ANS) binding	144
	5.4.2 Susceptibility to thermolytic proteolysis	145
5.5	'Self-chaperoning' of GroEL in the presence of (Mg-ATP)	148
5.6	Discussion	148

Chapter 6

General discussion and future prospects

6.1	Introduction	153
6.2	The ATP binding site of GroEL	153
6.3	GroEL assembly	154
6.4	Unanswered questions in the field of chaperone proteins	156
6.5	Evolution of the chaperonin family	157
6.6	Future prospects	159

References 161

Appendix: List of publications 181

List of figures

Short	title	Page no.
1.1	Model for the GroEL-GroES assisted folding of polypeptides	39
2.1	Plasmid pGT3270	47
3.1	SDS-PAGE analysis of GroEL eluting at Sephacryl S-200 void volume	67
3.2	Chromatography on Mono Q	68
3.3	Chromatography on Alkyl-Superose	69
3.4	Chromatography on Superose 6	70
3.5	SDS-PAGE analysis of purified GroEL	71
3.6	Molecular mass of GroEL	72
4.1	2', 3' dialdehyde ATP (oATP)	77
4.2	Kinetics of inactivation of GroEL ATPase with oATP	
	A. Pseudo first-order plots for inactivation	80
	B. Determination of K_I and k_2	80
4.3	Substrate protection against oATP inactivation of GroEL ATPase activity	
	A. Pseudo first-order plots for inactivation	83
	B. Protection against oATP inactivation by inclusion of ATP	83
4.4	Stoichiometry of inactivation of GroEL ATPase activity by oATP	86
4.5	SDS-PAGE analysis of thermolysin limited proteolysis of GroEL	90
4.6	Reverse phase HPLC profiles, monitored at 214 nm	91
4.7	Reverse phase HPLC profiles, monitored at 260 nm	93
4.8	Reverse phase HPLC profiles, monitored at 214 nm, of the modified	

	GroEL peptide	
	A. Rechromatography on μBondapak C18 column	95
	B. Chromatography on a microbore C8 column	95
4.9	Electrospray mass spectroscopy of oATP modified peptide	
	A. Electrospray in the positive ion mode	100
	B. Electrospray in the negative ion mode	100
4.10	Reaction of thiol with DTNB	103
4.11	Proteinase K limited proteolysis of GroEL	
	A. SDS-PAGE analysis of proteinase K digested GroEL	108
	B. Effect of limited proteolysis on the ATPase activity of GroEL	108
4.12	SDS-PAGE analysis of proteinase K treated-GroEL in the presence of oATP	110
4.13	Proposed reaction of oATP with active site cysteine	117
4.14	Sequence alignments of proposed ATP triphosphate binding regions in chaperonin 60 proteins	120
5.1A	Changes in the far-uv CD spectrum of GroEL in the presence of GdnHCl	125
5.1B	Changes in ellipticity at 225 nm in the far-uv CD spectra	127
5.2	Changes in the fluorescence of GroEL in the presence of GdnHCl	
	A. Excitation at 290 nm	129
	B. Excitation at 280 nm	129
	C. Changes in fluorescence at 315 nm	131
5.3	Changes in M_r of GroEL in the presence of GdnHCl	133
5.4	Changes in the ATPase activity of GroEL in the presence of GdnHCl	134
5.5	Changes in the ATPase activity and structural parameters of GroEL in the presence of GdnHCl	136
5.6	Reactivation of GroEL after incubation in GdnHCl	139
5.7	Changes in the far-uv CD spectra of GroEL after removal of GdnHCl by dialysis	140

5.8 Changes in the fluorescence of GroEL after removal of GdnHCl by dialysis

A. Excitation at 290 nm 142

B. Excitation at 280 nm 142

5.9 ANS fluorescence of GdnHCl denatured-GroEL 146

5.10 Susceptibility of GroEL to proteolysis by thermolysin 147

List of tables

Short	title	Page no.
1.1a	Sub-classes of chaperone proteins	3
1.1b	References for table 1.1a	4
1.2	Polypeptides which can be released from GroEL in the absence of ATP hydrolysis	37
4.1	Half life ($t_{1/2}$) and first order rate constants (k_{obs}) for the inactivation of GroEL ATPase activity at different oATP concentrations	78
4.2	Effect of substrate concentration on oATP inactivation of GroEL ATPase activity	82
4.3	Stoichiometry of inactivation of GroEL ATPase activity by oATP	87
4.4	Automated Edman degradation of the oATP modified material	98
4.5	Sulphydryl titration of GroEL	102
4.6	Automated Edman degradation of proteinase K treated-GroEL	105
4.7	Stoichiometry of modification of oATP into proteinase K-treated GroEL	106

Abbreviations

The abbreviations used in this thesis are set out in the Biochemical Journal "Instructions to Authors", except the following:

A	absorbance
Amp	ampicillin
ADP	adenosine diphosphate
ANS	8-anilino-1-naphthalenesulphonic acid
ATP	adenosine triphosphate
AUFS	absorbance units full scale
BSA	bovine serum albumin
cAMP	adenosine 3': 5'-cyclic monophosphate
CAPS	3-[cyclohexylamino]-1-propanesulphonic acid
CD	circular dichroism
Cpn	chaperonin
(k)Da	(kilo) daltons
DEAE	diethyl aminoethyl
DNAse	deoxyribonuclease
DTNB	5, 5'-dithio-bis (2-nitrobenzoic acid)
DTT	dithiothreitol
EDTA	ethylene diamine tetra acetate
E. R.	endoplasmic reticulum
FPLC	fast protein liquid chromatography
5'-p-FSBA	5'-p-fluorosulphonylbenzoyladenosine
GdnHCl	guanidinium chloride
HPLC	high pressure liquid chromatography
Hsc	heat shock cognate
Hsp	heat shock protein

Hy	hydrophobic amino acid
M_r	molecular weight
NADH	reduced nicotinamide adenine dinucleotide
N.M.R.	nuclear magnetic resonance
NTB	nitrothiobenzoate
oATP	2', 3' dialdehyde ATP
ODx	optical density at x nm measured with a 1 cm path
Pi	orthophosphate, inorganic phosphate
PAGE	polyacrylamide gel electrophoresis
PDI	protein disulphide isomerase
PEP	phospho*enol* pyruvate
PMSF	phenylmethanesulphonylfluoride
PPIase	peptidyl prolyl *cis-trans* isomerase
PVDF	polyvinyldiene difluoride
Rf	mobility of a protein compared to the dye front
rpm	revolutions per minute
Rubisco	ribulose 1, 5-*bis* phosphate carboxylase
SDS	sodium dodecyl sulphate
TCA	trichloroacetic acid
TEA	triethanolamine
TEMED	N, N, N', N'-tetramethylethylene diamine
TFA	trifluoroacetic acid
Tris	Tris (hydroxymethyl) aminomethane
UV	ultra violet
Xaa	any amino acid

ACKNOWLEDGEMENTS

I would like to thank the following people and organisations without whose help this work could not have been completed.

Prof. John Coggins and Dr. Nicholas Price for their discussions, help and supervision throughout this project. The Biochemistry Department, University of Glasgow, for providing the facilities necessary for this work. The Science and Engineering Research Council for the funding of this project. Dr. Neil McLennan at the University of Edinburgh for providing the over-producing *E. coli* strains used in this work. Ms. Sharon Kelly at the University of Stirling for her expertise in recording the CD spectra displayed in this thesis. Dr. Bryan Dunbar and Prof. John Fothergill at the University of Aberdeen for their assistance in peptide sequencing. Mr. John Greene for his invaluable technical assistance, Mr. Malcolm Horsburgh for proof-reading parts of this thesis and Dr. J. Ellbee for computing assistance. All colleagues, past and present, especially those in B4a and D-floor for their help and friendship over the past three years. In particular I would like to thank Tim, Will, Gillian, Fiona, Graeme, Ranjit, Gavin, Colin, Michele, Sally, Paul, Ritchard, Pam, Carol, Denis Bergkamp and Ronald Koeman. I would also like to acknowledge the following hostelries, Tennents, Uisge Beatha, the Brewery Tap and others which are too numerous to mention for providing, in someones words, a 'focus in life'. Finally I must thank my parents for their support and encouragement throughout my studies.

SUMMARY

The reaction of the *E.coli* chaperonin GroEL (cpn 60) with the ATP analogue 2', 3' oxidised ATP (oATP) was studied with a view to identify the important amino acid(s) present at the ATP binding site of GroEL. Treatment with the reagent leads to loss of the ATPase activity of GroEL in a pseudo-first order fashion; this can be prevented by inclusion of ATP in the reaction mixture. Measurements of the stoichiometry of the reaction indicate that the loss of activity corresponds to the incorporation of about one oATP per subunit of GroEL. From analysis of the sequences of modified peptides it is proposed that the reaction probably occurs with one or both of the two cysteines Cys 457 and Cys 518, although the instability of the adduct(s) makes a definite identification of the site(s) of reaction difficult. The involvement of Cys side chains in the reaction with oATP was confirmed by using DTNB (5,5'-dithiobis(2-nitrobenzoate)) to estimate thiol groups in both modified and unmodified GroEL.

The unfolding of GroEL in solutions in guanidinium chloride (GdnHCl) was also studied. From the results of CD, fluorescence and light scattering it is clear that major structural transitions in the protein occur over the range 1.0 - 1.5 M-GdnHCl, whilst the ATPase activity is lost at lower concentrations (0.75 M). The unfolding of the protein appears to be a highly cooperative process. After denaturation in concentrations of GdnHCl above 1.0 M, removal of the denaturing agent by dialysis results in the very nearly complete regain of secondary structure (as judged by CD) but not the regain of correct tertiary structure or quaternary structure, nor ATPase activity. The product was shown to be very sensitive to proteolysis by thermolysin, unlike the native protein, but did not show enhanced binding of ANS, a characteristic property of the 'molten globule' state of proteins. .

Chapter 1 Introduction

1.1 Molecular chaperones: general introduction

The first-half of the genetic pathway whereby information from linear genomic DNA is transferred into a three-dimensional functional protein has been well characterised, but until recently relatively little research has been undertaken to elucidate the steps the nascent polypeptide undergoes from emerging from the ribosome to folding into its final native structure. The classic work of Anfinsen (1973), on his studies on the renaturation of ribonuclease A *in vitro*, led to the proposal that many proteins are able to fold to the native state spontaneously and that the amino acid sequence of a protein contains all the relevant information to specify its three-dimensional structure. The extrapolation from the results of these types of experiments was that a protein *in vivo* could fold and assemble spontaneously in a process intrinsic to its primary structure, independent of other factors. Many *in vitro* refolding experiments, however, are successful only under non-physiological conditions *i.e.* at low temperatures (which attenuates hydrophobic interactions) and low protein concentrations (to limit aggregation). The relatively high protein concentrations and temperature of the cytosol can subject the growing polypeptide chain to premature inter- or intra-polypeptide interactions, often leading to misfolding and aggregation. Thus whilst some polypeptides, and especially small globular polypeptides, may fold spontaneously *in vitro* to attain native structure it is becoming increasingly obvious that others require the assistance of accessory proteins.

To deal with such problems a group of proteins known as chaperone proteins have evolved in order to ensure that polypeptides are able to fold or be transported properly. The term 'molecular chaperone' was first used by Laskey *et al.* (1978) to describe the function of nucleoplasmin. Nucleoplasmin binds to histones in order to shield their strong positive charge and as such prevents the misassembly of nucleosomes from separated histones and DNA. In the absence of nucleoplasmin, at physiological ionic strength, self-assembly fails due to the formation of a precipitate

between the DNA and histone monomers. Although nucleoplasmin is the archetypal chaperone protein the definition has been extended in order to comprise a variety of sub-classes of chaperone proteins (Table 1.1). It is likely that this list will grow longer as other novel chaperone proteins are identified. The major classes are the heat shock proteins 60, 70 and 90 which will be discussed later. The chaperone proteins discovered to date do not, however, convey any steric information towards the folding process. Instead they are believed to have a passive role in that they simply act to prevent any off-pathway folding processes such as aggregation. Therefore the 'self-assembly' hypothesis of Anfinsen is not violated as the final three-dimensional structure is indeed dictated by the primary amino acid sequence. Proteins which fold and assemble only in the presence of chaperone proteins may be seen to be undergoing a process of 'assisted self-assembly' (Ellis, 1990).

1.2 Protein folding *in vivo* versus protein folding *in vitro*

Ideally the studying of protein folding should be performed *in vivo*, with the translation of the mRNA synchronised to the folding of the protein. The opportunity exists for refolding to commence on the emergence of the nascent polypeptide from the ribosome. Evidence exists that disulphide bonds can be formed co-translationally, this has been observed in the folding of rat-serum albumin (Peters and Davidson, 1982) and of immunoglobulins (Bergman and Kuehl, 1979). The evidence that short peptides can indeed have ordered secondary structure, for example the N-terminal first 45 residues of adenylate kinase can still bind MgATP tightly like the native enzyme presumably due to having a conformation similar to that in the enzyme (Fry *et al.*, 1985), is consistent with the proposal that folding of proteins is initiated on the emergence of the polypeptide from the ribosome (Bergman and Kuehl, 1979). However protein folding is not entirely co-translational in that protein folding does not cease when translation ends, as newly synthesised polypeptides undergo constant adjustments in order to attain their final native structure. Such evidence for co-translational folding, however,

Table 1.1a Sub-classes of chaperone proteins

Class	Members	Proposed roles
1. Nucleoplasmin	Nucleoplasmin, Protein XLNO-38	Nucleosome assembly,
	Protein Ch-NO38, Nucleoplasmin S	Ribosome assembly & transport
2(a) Chaperonins	Chaperonin-60:	
	GroEL (*E. coli*), Hsp 60, mitonin, HuCha 60	Stress protection,
	(all mitochondrial), ch-cpn60, rubisco	Protein folding,
	subunit binding protein (both plastid)	Oligomer assembly,
	Chaperonin-10:	Protein transport,
	GroES (*E. coli*), hsp 10 (mitochondria),	DNA replication,
	cpn 10 (plastid)	mRNA turnover
2(b) TRiC family	TF55 (archaebacteria), TRiC (mammals)	
3. HSP 70	Hsp 68, 72, 73; DnaK; clathrin uncoating	Protein folding,
	ATPase;BiP; Grp 75, 78, 80; hsc70,	Oligomer assembly /
	KAR 2; SSC 1; SSD 1; SSA 1-4;	disassembly, Protein
	SSB 1	transport, Stress protection
4. HSP 90	Hsp 83, 87; Htp G; Grp 94	Masking of hormone receptor
5. Signal recognition particle		Protein transport
6. Subtilisin prosequence		Subtilisin folding
7. α-lytic factor prosequence		α–lytic protease folding
8. Ubiquitinated ribosomal proteins		Eukaryotic ribosome assembly
9. Trigger factor		Protein transport
10. Sec B protein		Protein transport
11. Pap D protein		*E. coli* pilus assembly
12.sHSPs / α-crystallins		Protein folding / Stress protect.
13.Carboxypeptidase Y prosequence		Carboxypeptidase Y folding

Table 1.1b References for table 1.1a

	Class	References
1.	Nucleoplasmin	Laskey *et al.* (1978), Dingwall and Laskey (1990)
2(a)	Chaperonins	see Hallberg (1990), Ellis (1991), Gething and Sambrook (1992), Hendrick and Hartl (1993), Neupert and Pfanner (1993), section 1.5.6
2(b)	TRiC family	Trent *et al.* (1991), Lewis *et al.* (1992), Yaffe *et al.* (1992), Mummert *et al.* (1993)
3.	HSP 70	see Gething and Sambrook (1992), Hendrick and Hartl (1993), Neupert and Pfanner (1993), Zylicz (1993), section 1.5.2
4.	HSP 90	Sanchez *et al.* (1987), Borkovich *et al.* (1989), Picard *et al.* (1990), Nadeau *et al.* (1993)
5.	Signal recognition particle	Georgopoulos (1992)
6.	Subtilisin prosequence	Zhu *et al.* (1989), Shinde *et al.* (1993)
7.	α–lytic factor prosequence	Silen and Agard (1989), Baker *et al.* (1992)
8.	Ubiquitinated ribosmal proteins	Finley *et al.* (1989)
9.	Trigger factor	Lecker *et al.* (1989)
10.	Sec B protein	Lecker *et al.* (1989), Georgopoulos (1992)
11.	Pap D protein	Holmgren and Bränden (1989)
12.	sHSPs / α–crystallins	deJong *et al.* (1988), Horwitz *et al.* (1992) Jakob *et al.* (1993), Merck *et al.* (1993)
13.	Carboxypeptidase Y prosequence	Winther and Sørensen (1991)

has led to the proposal that the *in vitro* method of studying folding, *i.e.* starting with a completely unfolded polypeptide chain via chemical denaturation, is not a good model for the studying protein folding (Bergman and Kuehl, 1979; Tsou, 1988), as it is not comparable with the *in vivo* situation.

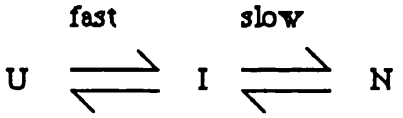
To characterise the conformational changes a polypeptide undergoes in reaching native structure *in vitro* studies are necessary. The *in vivo* versus *in vitro* problem described above refers mainly to the kinetics of (re)folding. Protein folding within the cell takes of the order of minutes (Tsou, 1988). Refolding *in vitro* varies in time from seconds, monomeric fructose 1, 6, bisphosphate aldolase can regain activity within 10 seconds after denaturation in guanidinium chloride (Rudolph *et al.*, 1983), to hours and is often dependent on the experimental conditions chosen. While the latter time scale is incompatible with protein folding *in vivo*, it is still much more rapid than protein folding via a random pathway. Levinthal (1968) has estimated that it would take of the order of 10^{50} years for a polypeptide of 100 amino acids to attain a particular conformation by random search. Part of the discrepancy between *in vivo* and *in vitro* (re)folding may be accounted by the fact that the cell contains additional 'tools' for efficient refolding, such as protein disulphide isomerase and peptidyl prolyl *cis-trans* isomerase which can accelerate the rate of protein folding. These proteins will be discussed in a later section. Despite any discrepancies between the time-scale of (re)folding, *in vitro* conformations, biological activity and relative molecular mass of a refolded protein are the same as would be found within the cell (Jaenicke, 1987). This would suggest that the *in vitro* method of studying protein folding, by denaturation with a strong denaturant followed by its removal to initiate the refolding process, is indeed a valid way of providing valuable information about the refolding process.

The efficiency of folding *in vitro* is very variable. While small monomeric proteins, especially those which lack disulphide bridges, can refold *in vitro* at rates comparable with the *in vivo* system some more complex proteins are unable to regain

activity at all after denaturation. This has been noted especially for large multimeric proteins such as the hexameric glutamate dehydrognase (West and Price, 1988) and ribulose 1, 6, bisphosphate carboxylase (Rubisco)(Goloubinoff *et al.*, 1989). The efficiency of folding *in vivo* has been estimated to be greater than 95% of all newly synthesised polypeptides eventually attaining their correct three dimensional structure (Gething and Sambrook, 1992). However misassembly and misfolding clearly occur in the cell (and *in vitro*) under unbalanced physiological conditions (Pelham, 1986) and the cell must have a mechanism to compensate for this.

1.3 The protein folding problem

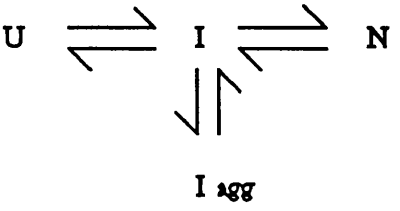
The isomerisation of proteins from the unfolded state to the native conformation involves the transient formation of folding intermediates. A general folding pathway for monomeric proteins may be described in simple terms by the equation



where U, I and N represent the unfolded, intermediate and native states, respectively. The initial fast steps in protein folding involve the rapid formation of secondary structure, which occur on the millisecond time scale, and the formation of a 'molten globule' type structure, as denoted by I in the above equation (Ptitsyn et al., 1990; Christensen and Pain, 1991). The 'molten globule' is believed to be a general intermediate in protein folding, and is characterised by having a compact yet mobile structure that lacks the close packing of the secondary structure elements of the native protein and due to the fluidity of the structure the amino acid side chains are more accessible to solvent molecules than in the folded protein. The 'molten globule' form exists in rapid equilibrium with the unfolded state. The rate determining step in protein

folding is, however, the slower, cooperative formation of specific tertiary structure associated with the native structure. It is the merging of pre-existing units from the semi-ordered 'molten globule' to the final adjustments in tertiary structure which are the limiting rate for a protein to attain native structure. These slow protein folding steps may involve disulphide-bond arrangements, isomerisation of X-Pro peptide bonds or domain pairing in the case of multi-domain proteins. In the case of the first two examples the rate-determining step in protein folding can be accelerated by the enzymes protein disulphide isomerase and peptidyl prolyl *cis-trans* isomerase, respectively (section 1.4). The folding of multimeric proteins is more complex and has been the subject of a number of reviews (see Price, 1992) and for reasons of brevity will not be discussed in this section .

Incorrect interactions can often take place during protein (re)folding leading to the formation of inactive aggregates due to incorrect inter- or intra-polypeptide interactions. This has been observed both *in vitro* (see section 1.5) and *in vivo*, in the case of formation of inclusion bodies during protein over-expression (Lee and Olins, 1992; Edgerton *et al.*, 1993) or during stress conditions within the cell (Pelham, 1986; Nguyen *et al.*, 1989). The 'molten globule' states of proteins are less soluble than the native protein, presumably due to having a greater number of hydrophobic residues exposed and as such have a penchant for aggregation. The kinetic partitioning of the I state between the N state and the aggregated state (I_{agg}) is therefore of great significance for the efficient folding of proteins.



In vitro one can control some of the factors which lead to aggregation. The simplest ways to suppress aggregation during *in vitro* refolding are by reducing the concentration of the protein, in order to reduce the amount of potential intra-molecular interactions, or by reducing the temperature of the refolding reaction, which suppresses hydrophobic interactions. In the complex cellular milieu such approaches are obviously not feasible and in order to resolve this dilemma the cell has apparently evolved molecular chaperones in order to modulate *in vivo* folding reactions (section 1.5).

1.4 Catalysis of protein folding

In vitro (and thus presumably *in vivo*) the rate determining steps of protein folding involving the isomerisation of covalent bonds can be catalysed by the enzymes described below. These enzymes, protein disulphide isomerase and peptidyl prolyl *cis-trans* isomerase, do not determine the folding pathway but act simply so as to accelerate the slow rate-determining steps of protein folding.

1.4.1 Protein disulphide isomerase

Protein disulphide isomerase (PDI) is an abundant protein found within the lumen of the endoplasmic reticulum (E.R.), which catalyses the thiol : disulphide interchange reactions leading to protein disulphide bond formation, isomerisation or reduction (depending on the polypeptide substrate and the imposed redox potential). The level of secretory protein synthesis in the E.R. of different cell types correlates with the abundance of PDI (Freedman, 1989). *In vitro* PDI can accelerate the formation of native interchain disulphide bonds in both type I and type II procollagen (Koivu and Myllylä, 1987) whilst co-translational studies have shown that the formation of disulphide bonds in γ-gliadin, in dog pancreatic microsomes, is dependent on the presence of PDI (Bulleid and Freedman, 1988). *In vivo* crosslinking studies have also

demonstrated that there is a specific association between newly synthesised immunoglobulin chains and the enzyme (Roth and Pierce, 1987).

The mammalian enzyme is a homodimer with identical subunits of M_r 57,000 each containing duplications of domains which show a strong homology to thioredoxin (Edman *et al.*, 1985). The active site of PDI, like thioredoxin, contains a pair of conserved vicinal cysteine residues in the sequence WCGHCK which participate in the thiol : disulphide exchange reactions. PDI has also been shown to be highly homologous to a number of other proteins suggesting that it may have multiple roles concerned with protein modification and assembly within the E.R.. The enzyme is identical to the β-subunit of prolyl-4-hydroxylase (Pihlajaniemi *et al.*, 1987), an enzyme which causes extensive modification of proline residues in nascent collagen molecules, although it is unknown, as yet, whether the active site cysteine residues involved in 'thiol shuffling' are also involved in the hydroxylation reaction. In addition, PDI is also a component of the microsomal triglyceride transfer protein complex which facilitates the incorporation of triglycerides into nascent very low density lipoproteins within the E.R.(Wetterau *et al.*, 1990). PDI has been shown to be related, on a sequence level, to a number of other proteins within the cell including a glycosylation site binding protein and thyroid hormone binding protein (Freedman, 1989).

Intriguingly PDI has also been proposed to be a chaperone protein, albeit one which also contains PDI activity, due to its multifunctionality and its general non-specificity towards substrate polypeptides (Wang and Tsou, 1993). The β-subunit of prolyl-4-hydroxylase (*i.e.* PDI) is essential for the correct folding and assembly of the $\alpha_2\beta_2$ functional complex of prolyl-4-hydroxylase, as it prevents the misfolding and aggregation of the α–subunit (John *et al.*, 1993). This may be considered a chaperone-like function although, unlike other known chaperone proteins, the β-subunit is also part of the final functional complex. Consistent with the proposal that PDI may also have chaperone-like activity the isomerase activity of PDI can be inhibited by the

alkylation of the conserved vicinal cysteine residues but the modified enzyme can still bind to peptides (Noiva *et al.*, 1991). The enzyme can apparently bind peptides without having any obvious sequence specificity ; this is similar to chaperone proteins which also show no obvious sequence specificity (discussed later). The four thioredoxin-like domains of mammalian PDI account for only approximately 40% of the entire polypeptide (Edman *et al.*, 1985). The thioredoxin-like domains may be responsible for the isomerase activity, whist the remainder of the chain may be responsible for non-specific peptide binding, that is the chaperone-like activity. Analogously, the peptide binding region of PDI has recently been located near the C-terminus, distinct from the PDI active site (Noiva *et al.*, 1993). Interestingly a recent paper has provided more direct evidence of this putative chaperone-like role of PDI. PDI is able to bind to a variety of denatured polypeptides and, along with other proposed E.R. chaperone proteins including BiP (a known E.R. chaperone protein), is released only in the presence of ATP (Nigam *et al.*, 1994). It remains for this interesting proposal to be confirmed.

1.4.2 Peptidyl prolyl *cis-trans* isomerase

Peptidyl prolyl *cis-trans* isomerase (PPIase) catalyses one of the rate-limiting steps in *in vitro* protein refolding, namely the slow cis-trans isomerisation of proline peptide bonds. The efficiency of catalysis by PPIase is variable and dependent on the target protein. The slow steps of refolding can be accelerated rapidly in the case of immunoglobulin light chains (Lang *et al.*, 1987) and ribonuclease T1 (Fischer *et al.*, 1989) whilst more moderate catalysis is observed with porcine ribonuclease (Lang *et al.*, 1987) and type III collagen (Bächinger, 1987). In other cases such as in the refolding of bovine ribonuclease A, no acceleration in the slow step of refolding is observed in the presence of PPIase, although the functionally distinct protein disulphide isomerase can catalyse more efficient refolding (Lang and Schmid, 1988). The isomerisation of Pro-peptide bonds is a relatively late step in protein (re)folding. Lang

et al. (1987) have proposed that the variability in catalysis of prolyl isomerisations may be due to the fact that in some cases the Xaa-Pro bond is simply inaccessible due to the protein having reached some degree of 'nativeness'

Proteins with PPIase activity are highly abundant and widely distributed and found in all tissues and organisms from bacteria to mammals. In addition, two structurally unrelated classes of PPIases exist, characterised by their inhibitor and substrate specificities. The first class of PPIase is identical to cyclophilin, a protein known to bind to the immunosuppressive drug cyclosporin A, (Takahashi *et al.*, 1989; Fischer *et al.*, 1989) whilst the other class of PPIase bind the structurally distinct immunosuppressant FK506 (Siekierka *et al.*, 1989; Harding *et al.*, 1989). Whilst the role of PPIase in immunosuppression warrants close attention the main *in vivo* role of PPIases is, however, believed to be in the catalysis in the rate of cis-trans isomerisation of Xaa-Pro peptide bonds during protein folding although, as yet, only circumstantial evidence for this *in vivo* role exists. Mutant flies with a defect in the *ninA* gene, which codes for a homologue of cyclophilin-PPIase, show a 10 fold reduction in the levels of rhodopsin although the mRNA is produced normally. The *ninA* gene product may therefore be necessary for the folding and stability of rhodopsin (Shieh *et al.*, 1989; Schneuwly *et al.*, 1989). In addition the *in vivo* folding of the triple helix of type I collagen in chick fibroblasts is delayed in the presence of cyclosporin A (Bächinger and Compton, 1991). More direct evidence is required to fully implicate PPIase as a catalyst in cellular protein folding. It has also been recently proposed that PPIase may, as has been suggested for PDI, also have a chaperone-like activity, for example in preventing the aggregation of carbonic anhydrase (Freskgard *et al.*, 1992).

1.5 Molecular chaperone proteins

Molecular chaperones have been defined as 'a family of unrelated classes of protein that mediate the correct assembly of other polypeptides but are not themselves

components of the final functional structure' (Ellis and van der Vies, 1991). The term mediate was deliberately chosen because the exact mechanism of chaperone mediated protein folding is still not yet fully understood. Chaperone proteins discovered to date do not act as catalysts of protein (re)folding, the recent proposals that PDI and PPIase may have chaperone-like activity (section 1.4) notwithstanding, but rather act to prevent off-pathway processes such as aggregation (section 1.3). The various chaperone proteins (table 1.1) are structurally homologous with each other, within each class, but not between classes. The major chaperone proteins, with particular emphasis on GroEL, will be discussed in this section.

1.5.1 The Nucleoplasmins

Nucleoplasmin is an acidic soluble nuclear protein which is the most abundant protein in amphibian oocytes, constituting between 7 and 10% of the actual nuclear protein in *Xenopus* oocytes (Mills *et al.*, 1980). The chaperoning function of nucleoplasmin was first described by Laskey *et al.* (1978) who implicated the protein in the formation of nucleosome cores from separated histones and DNA in *Xenopus* oocyte extracts. At physiological ionic strength nucleosome cores are unable to form, in the absence of nucleoplasmin, due to the formation of a precipitate between the basic histones and acidic DNA. In the presence of a molar excess of nucleoplasmin the strong positive charge of the histones is shielded by the binding of the chaperone, and thus initial non-specific DNA-histone interactions are prevented allowing the formation of the most energetically stable conformation, the nucleosome. Nucleoplasmin assisted nucleosome assembly is believed to be Mg-ATP independent (Laskey *et al.*, 1978), which distinguishes it from other known chaperone proteins such as the Hsp 70 and chaperonin 60 class. Under non-physiological conditions nucleosomes can also be reconstituted from separated DNA and histones, in the absence of nucleoplasmin, by prolonged dialysis from high salt concentrations (2 M-NaCl) to lower salt

concentrations (Felsenfeld, 1978), indicating that the steric information for nucleosome assembly resides in the histones and not in nucleoplasmin.

Since the discovery of nucleoplasmin several homologues have been discovered, including the nucleolar protein XLNO-38 which may mediate the assembly of ribosomes. This suggests that a family of nucleoplasmin-like proteins exists, the members of which carry out various chaperoning functions within the cell (Dingwall and Laskey, 1990).

1.5.2 The Hsp70 family

1.5.2.1 General background

The heat shock 70 class (Hsp 70) class of molecular chaperones is an abundant family of stress proteins which are found throughout eukaryotic cells. In yeast members include the cytosolic proteins Ssal-4p, endoplasmic reticulum Kar2p and the mitochondrial protein Ssc1p; homologues in mammals identified to date include the cytosolic Hsc73 protein and BiP (or Grp78) which is found in the endoplasmic reticulum. Members of this family have also been identified in plastids, namely ct-Hsp70. These proteins have been conserved throughout evolution, indeed the Hsp70 homologue in *E. coli* , DnaK, shows approximately 50% sequence identity at the amino acid level to Hsp70 of eukaryotes (Bardwell and Craig, 1984). These proteins were initially discovered as stress proteins which were induced during the exposure of the cell to heat or other forms of stress. Under such conditions polypeptides can become partially denatured, exposing hydrophobic regions, and thus are prone to aggregation. This led to the proposal that Hsp70 can interact with such proteins in order to limit any such aggregation and, perhaps, promote disaggregation (Pelham, 1986). However, not all members of the Hsp70 proteins are heat-shock inducible, some are constitutively expressed and thus it has become clear that these proteins also have essential functions

in the normal cell, particularly in *de novo* protein folding. These proteins have been shown both *in vitro* and *in vivo* to bind and stabilise unfolded forms of polypeptides in order to facilitate membrane translocation and / or their subsequent proper folding and assembly (for recent reviews see Craig and Goss, 1991; Gething and Sambrook, 1992; Hendrick and Hartl, 1993).

1.5.2.2 Conserved structure of the Hsp70 class

The Hsp70 proteins bind ATP and have a weak ATPase activity which has been implicated in their function. Comparison of the amino acid sequences of known Hsp70 proteins show that the N-terminal two-thirds (~450 a.a.'s) are more highly conserved than the C-terminal region. Protease treatment of bovine clathrin uncoating ATPase (Hsc70) has identified an amino terminal 44-kDa resistant fragment that retains ATPase activity which is uncoupled from clathrin binding (Chappell *et al.*, 1987). The structure of this domain has been determined by X-ray diffraction to a resolution of 2.2 Å (Flaherty *et al.*, 1990), revealing two domains of similar fold on either side of a large cleft with an ATP binding site at the bottom of the cleft. Despite the differences in function and amino acid sequence of the proteins the three-dimensional arrangement of the 44-kDa fragment of Hsc70 is similar to that of both rabbit skeletal muscle actin (Flaherty *et al.*, 1991) and the nucleotide binding site of yeast hexokinase (Flaherty *et al.*, 1990). The ATP binding motif of these type of proteins - proposed to have evolved from a common ancestral origin (Bork *et al.*, 1992) - is, however, distinct from other ATP binding proteins such as recA protein and adenylate kinase which contain a type A phosphate binding motif as described by Walker *et al.* (1982). This latter type of nucleotide-binding protein themselves have been proposed to have a common ancestral core (Milner-White *et al.*, 1991) and, may, have evolved independently from the ATP binding protein class of which Hsc70 is a member.

The variable C-terminal region of the Hsp70 proteins is responsible for polypeptide binding (Chappell *et al.*, 1987) and has been identified as an 18-kDa region of Hsc70 protein located immediately after the 44-kDa ATPase domain (Wang *et al.*, 1993). The three-dimensional structure of the peptide binding domain is not known and several investigators have attempted to define this region, particularly in view of the rather catholic nature these chaperone proteins have towards 'substrate' polypeptides. This region of the protein has, however, been predicted to resemble the $\alpha_1\alpha_2$ superdomain (*i.e.* the peptide binding site) of the human MHC class 1 antigen HLA based on a consensus secondary structure deduced from the amino acid sequences of 33 Hsp70 proteins (Rippmann *et al.*, 1991). Data would seem to indicate that Hsp70 chaperone proteins interact with short stretches of (mainly) hydrophobic amino acids (Flynn *et al.*, 1991; Richarme and Kohiyama, 1993; Blond-Elgundi *et al.*, 1993a; Gragerov *et al.*, 1994) which exist in an extended conformation (Landry *et al.*, 1992; Blond-Elgundi *et al.*, 1993a). Taking advantage of the fact that binding of some, but not all, short synthetic peptides stimulates Hsp70 ATPase activity, Flynn *et al.* (1991) synthesised a series of peptides of random amino acid sequence of known chain length in order to study binding to BiP, the mammalian endoplasmic reticulum Hsp70. These authors found that the peptide binding site of BiP is filled by a stretch of 7 residues, which are enriched in aliphatic amino acids, particularly in positions 3-6. Taking this a step further Blond-Elgundi *et al.* (1993a) used a bacteriophage display system to search many peptide sequences which could be recognised by BiP. This led to the discovery of a heptapeptide motif which has been implicated in recognition by BiP, namely Hy-(W/X)-Hy-X-Hy -X-Hy, where Hy is a large hydrophobic amino acid (most frequently Trp, Leu or Phe), W is Trp and X is any amino acid. The (poly)peptide binding site would therefore be seen to contain 4 pockets that can accommodate the large side chains of these hydrophobic and aromatic amino acids. Interestingly this type of arrangement is similar to that described for the binding of peptides to major histocompatability antigens (see Brown *et al.*, 1993). The *E. coli* Hsp70 homologue DnaK also preferentially interacts with short stretches of hydrophobic sequences (Richarme and

Kohiyama, 1993; Gragerov *et al.*, 1994) although, as yet, no general consensus sequence has been identified. Such extended hydrophobic stretches are likely to be exposed only during the early stages of polypeptide synthesis and / or folding, presumably being buried in the native protein. Such hydrophobic peptides may be targets for Hsp70 chaperone proteins which could thus prevent off-pathway interactions occurring *in vivo* between the exposed hydrophobic regions which would otherwise limit translocation and / or correct folding. The fact that Hsp70 proteins recognise such target motifs in an extended form (Landry *et al.*, 1992) which contain little or no stable secondary structure, such as reduced carboxymethylated α-lactalbumin (Langer *et al.*, 1992a), distinguishes them from the chaperonin 60 (Hsp60) class of chaperone proteins which seem to recognise secondary structure elements. This is likely to account for the differing roles of the Hsp70 and the chaperonin 60 class within the cell and will be discussed in a later section (section 1.6).

1.5.2.3 General mechanism for Hsp70 chaperone mediated action

Hsp70 proteins, certainly in the case of the E. R. protein BiP, are believed to exist in a oligomeric form which can be mobilised into an active monomeric form on the binding of substrate polypeptides (Blond-Elguindi *et al.*, 1993b). They interact with unfolded segments of the polypeptide chain - *i.e.* a nascent polypeptide emerging from the ribosome, or after membrane translocation through the lipid bilayer or with sequences exposed by partial denaturation after cellular stress. The interaction is via their C-terminal region. The binding of the chain prevents any (further) premature inter- or intra-polypeptide interactions from occurring which may lead to aggregation or non-productive folding. The binding of Hsp70 proteins to both protein and peptide substrates is weakened in the presence of hydrolysable ATP (Flynn *et al.*, 1989; Palleros *et al.*, 1991; Liberek *et al.*, 1991b; Landry *et al.*, 1992; Langer *et al.*, 1992a). Nucleotide binding to the N -terminus causes a conformational change in the Hsp70 protein, as judged by protease susceptibility (Liberek *et al.*, 1991b), which is

presumably transmitted to the C-terminal domain, whereupon the polypeptide is released. The released polypeptide may then fold into its native state, as a monomer or oligomer, or interact with other members of the chaperone machinery (Section 1.6) before attaining native structure. *In vivo*, however, the Hsp70 chaperones are believed to interact with other proteins which help to regulate the chaperones ATPase activity and thus polypeptide release (section 1.5.2.4).

1.5.2.4 Regulation of Hsp70 function

The mechanism of Hsp70 mediated action is not quite as straightforward as described above, due to the discovery of a number of accessory proteins which regulate the (ATPase) activity of these chaperone proteins. The *E. coli* Hsp70 homologue DnaK is regulated by the 41-kDa chaperone DnaJ and the 20-kDa GrpE, a constitutively expressed stress protein. The *dna* K gene was initially identified along with *dna* J and *grp* E as *E. coli* genes required for the growth of bacteriophage λ (Friedman *et al.*, 1984) and all three gene products have since been implicated as being important for normal cellular function.

Although a molar excess of DnaK on its own, in the presence of ATP, can prevent the thermal inactivation of RNA polymerase *in vitro* and reactivate it (Skowyra *et al.*, 1990), the presence of the other regulatory components is required, in other cases, to potentiate the effect of DnaK-mediated folding (Langer *et al.*, 1992a; Schröder *et al.*, 1993). *In vivo* and *in vitro* all three components of the DnaK, DnaJ and GrpE chaperone system, plus ATP, are required for the efficient repair of thermally inactivated luciferase (Schröder *et al.*, 1993). DnaJ itself is a chaperone protein, it can suppress the aggregation of luciferase, prior to the addition of DnaK and GrpE, and is known to bind to unfolded polypeptides such as rhodanese (Langer *et al.*, 1992a). In addition DnaJ recognises and binds nascent ribosome-bound polypeptides as short as 55 amino acids (Hendrick *et al.*, 1993). The role of DnaJ *in vivo* is believed to be to

protect the nascent polypeptide from aggregation then target the unfolded polypeptide to DnaK. This leads to the formation of a ternary complex between DnaJ, DnaK and polypeptide of which DnaK, stabilised in its ADP-bound state (via ATP hydrolysis on DnaJ / polypeptide binding), has a high affinity for 'substrate' polypeptides (Palleros *et al.*, 1993). The ATPase activity of DnaK is regulated by both DnaJ and GrpE, whereby DnaJ directly stimulates ATP hydrolysis and GrpE acts as a nucleotide exchange factor (Liberek *et al.*, 1991a). Release of the polypeptide from DnaK is via an ATP induced conformational change in this chaperone. This is mediated by GrpE which, via its nucleotide exchange activity, stimulates the release of ADP followed by the binding of ATP. It is believed that it is the binding of ATP, rather than its hydrolysis, which is required for polypeptide release (Palleros *et al.*, 1993). The ATP bound form of DnaK has a lower affinity for 'substrate' polypeptides and as such the 'chaperone complex' dissociates allowing release of the polypeptide to fold productively or be withdrawn by binding to another chaperone member (section 1.6). The polypeptide may, however, have to undergo several cycles of binding and release to DnaK, tightly regulated by DnaJ and GrpE, before reaching a state whereby it is committed to folding.

Several homologues of these regulatory proteins have also been found throughout eukaryotic cells. In yeast, DnaJ homologues include the cytosolic protein Ydj1p, the mitochondrial protein Scj1p and Sec63p, located in the endoplasmic reticulum. These proteins share several conserved features with DnaJ such as a cysteine-rich motif of unknown function and a putative domain, the 'J-region', which may mediate interactions with Hsp70s (Silver and Way, 1993). The mechanism of action described for the DnaK system is, therefore, also likely to be universal in the eukaryotic system..

1.5.3 The Hsp90 family

Like Hsp70s, members of the Hsp90 class of chaperones are highly conserved in bacteria, yeasts and mammals with approximately 40% identity between the various eukaryotic Hsp90s and the *E. coli* homologue HtpG (Bardwell and Craig, 1988). Hsp90s are highly abundant proteins, localised predominantly in the cytoplasm. In *S. cerevisiae* two homologues exist, Hsc82 - a constitutively expressed protein - and Hsp82, which has a low basal level of expression and which can be induced 10 to 15 fold on heat-shock (Borkovich *et al.*, 1989). Inactivation of both of these genes is lethal at any temperature, although cells can survive single mutations at temperatures up to 37.5°C.

The vertebrate Hsp90 proteins are found associated with a variety of proteins including cellular protein kinases and steroid hormone receptors (Gething and Sambrook, 1992) where the common role of the Hsp90 proteins, in these cases, seems to be in stabilising the target protein in an inactive or unassembled state. Hsp90-bound steroid receptors retain a high affinity for the steroid hormone but are unable to bind DNA and therefore unable to activate transcription. The displacement of Hsp90 from the receptor occurs on steroid-binding which then enables the steroid receptor to bind DNA and activate transcription (Sanchez *et al.*, 1987). This is not simply a case of steric hindrance by the Hsp90 protein. Glucocorticoid and estrogen receptors are unable to activate transcription in *S. cerevisiae* strains depleted in Hsp90 (Picard *et al.*, 1990) whilst *in vitro* studies have demonstrated that Hsp90 binding to glucocorticoid receptors is a prerequisite in order to maintain the hormone binding conformation of the receptor and, in turn, induce DNA-binding (Bresnick *et al.*, 1989). The role of the Hsp90 protein is, therefore, to interact with aporeceptors in order to convert them into an activatable conformation, whereby the hormone binding domain is established, and to inactivate any other functional domains which is reversed upon steroid binding (and Hsp90 release) (Picard *et al.*, 1990).

Further evidence of the chaperoning properties of Hsp90 comes from studies on chemically denatured proteins. Bovine pancreatic Hsp90, in a manner analogous to the *E. coli* chaperonin GroEL, suppresses the aggregation of citrate synthase and increases the yield of reactivation of both citrate synthase and the F_{ab} fragment of a monoclonal antibody (Wiech *et al.*, 1992). The binding to the target polypeptide is at a stoichiometry of 1 Hsp90 dimer per 1 or 2 molecules of unfolded polypeptide. In these cases there was no requirement for nucleoside triphosphates and it is therefore difficult to evaluate the significance for this role *in vivo* , in the absence of controlled binding and release. However it should be noted that the the F_{ab} fragment can also be released from the GroEL chaperone protein in the absence of ATP (Schmidt and Buchner, 1992) which may reflect the weak binding of this protein. Recently members of the Hsp90 class have been shown to have an ATPase activity, which is more potent than that of the Hsp70 chaperones, and undergo conformational changes upon ATP binding (Nadeau *et al.*, 1993). This may suggest that these proteins have a similar mechanism of action as the Hsp60 and Hsp70 proteins whereby the ATPase activity is regulated *in vivo* by accessory proteins. Such proteins, if they exist, have not been identified although Hsp90 have been shown to be associated with Hsp70 and both classes of peptidyl prolyl *cis-trans* isomerases (Nadeau *et al.*, 1993). It remains to be seen whether these proteins interact together to refold proteins in the heat-shock response.

1.5.4 The sHSP family

Small heat shock proteins (sHSP) are a ubiquitous and conserved group of heat-shock inducible proteins with a molecular mass of 15-30 kDa which exist in high molecular weight complexes within the cell. The sHSPs are homologous to α-crystallin (a predominant eye lens protein which is also present in other tissues) at both the amino acid sequence level (deJong *et al.*, 1988) and on a structural level (Merck *et al.*, 1993). The α-crystallins and the sHsps also share functional chaperone-like properties in

preventing misfolding of a variety of thermally and chemically denatured proteins. The α-crystallins suppress the thermally induced aggregation of a variety of proteins including the other major eye lens structural proteins β–crystallin and γ–crystallin and can assist the refolding of chemically denatured γ–crystallin (Horwitz, 1992). α-crystallin and sHsps have comparable functions in influencing the folding of other (non-eye lens specific) proteins under heat-shock conditions (Jakob *et al.*, 1993; Merck *et al.*, 1993). Interestingly the reactivation of both citrate synthase and α-glucosidase in the presence of sHsps or α-crystallin seem to be independent of ATP (Jakob *et al.*, 1993), similar to that described for the Hsp90 class (above). The mechanism of chaperoning-action of both sHsps and Hsp90 remains to be elucidated. It remains to be seen whether the sHSP (and possibly the Hsp90) class(es) have specialised functions within the cell, independent from the other chaperone proteins, or whether they act in concert with regulatory proteins *in vivo* .

1.5.5 Co-translational chaperones

The assembly of a number of polypeptides, which are synthesised as precursors, require the presence of their N-terminal pro-sequences in order to attain native structure. The correct (re)folding of subtilisin (Zhu *et al.*, 1989), α-lytic protease (Silen and Agard, 1989) and carboxypeptidase Y (Winther and Sørensen, 1991) is dependent on the presence of the respective pro-sequences in order to attain their final functional structures. These extra sequences assist the correct assembly of their target polypeptides before being cleaved. These pro-sequences have been termed 'co-translational chaperones' (Ellis and van der Vies, 1991) because they are covalently attached to the molecules whose assembly they mediate, thus having the advantage of interacting with their ligand from (potentially) the onset of translation without having to seek out their target in the complex cellular milieu. It should be noted that the co-translational chaperones are not strictly speaking a family of chaperones, *i.e.* no sequence homology exists between them, but rather they are grouped together because

of their similar functions. The mechanism by which co-translational chaperones assist polypeptide folding is, as yet, unknown. In the case of chemically denatured α–lytic protease, addition of the pro-sequence is required to overcome a kinetic block in the folding pathway to allow a catalytically inactive but folding competent state (resembling a 'molten globule') to attain native structure (Baker *et al.*, 1992). This is more in the manner of a folding catalyst rather than a conventional chaperone protein. No similar intermediates are present during the refolding of subtilisin. Shinde *et al.* (1993) have proposed that the pro-sequence of subtilisin is actually required from the onset to induce a folding competent state of the protein, characterised by the presence of residual secondary structure, and assist in formation of final native structure. This latter suggestion is more in line with the conventional definition of molecular chaperones and is possibly nearer what happens *in vivo* .

It should be noted that the proteins described above do not require the assistance of other chaperone proteins in order to fold correctly and as such differ from those proteins whose pro-sequences are insufficient to guide them to their final native structure in the absence of other chaperones.

1.5.6 The Hsp60 family:- the chaperonins

1.5.6.1 General history

The *Escherichia coli* protein cpn60 (GroEL) and its homologues in mitochondria (Hsp60) and chloroplasts (rubisco subunit binding protein or ch-cpn60) belong to a class of evolutionary related stress proteins which are members of the sub-class of chaperone proteins known as chaperonins (Hemmingsen *et al.*, 1988). GroEL is a heat shock protein which can be increased from about 1% to 10% of the total soluble cytoplasmic protein following cellular stress (Hemmingsen *et al.*, 1988) but is essential for cellular growth at all temperatures (Fayet *et al.*, 1989). The *groE* gene,

encoding GroEL and its partner chaperonin GroES, was initially discovered because mutations in it prevented the growth of several bacteriophages. This block is at the level of assembly of the bacteriophage where GroEL is required for the assembly of bacteriophage λ and T4 head structures and T5 tail structure (Georgopoulos *et al.*, 1972; Georgopoulos *et al.*, 1973; Zweig and Cummings, 1973). Rubisco subunit binding protein was discovered several years later associated with newly translated large subunits of rubisco in the stroma of *P. sativum* chloroplasts but absent in the mature protein (Barraclough and Ellis, 1980), whilst Hsp60 in mitochondria was first identified in *Tetrahymena thermophilia* as a GroEL related stress protein during heat shock conditions (McMullen and Hallberg, 1987). These early studies indicated a role for the chaperonin proteins in the assembly of other proteins and it now seems clear that this is a general function for these proteins within the cell. Highly-related Hsp60 homologues have not been detected in the eukaryotic cytosol or the endoplasmic reticulum. Proteins with weak Hsp60 homology have, however, been identified in the cytosol of thermophilic archaebacteria, plants and mammals which appear to possess chaperonin-like functions (Trent *et al.*, 1991; Lewis *et al.*, 1992; Yaffe *et al.*, 1992; Mummert *et al.*, 1993). These proteins may be regarded as a sub-class of chaperonins, namely the TRiC family, which are postulated to carry out GroEL-like functions within the eukaryotic cytoplasm.

1.5.6.2 The *gro* E gene products

GroEL is an oligomeric protein consisting of 14 identical subunits, with a subunit M_r of 57, 259 as deduced from its amino acid sequence (Hemmingsen *et al.*, 1988). Electron microscope analysis of negatively stained oligomeric GroEL indicates that the complex has an overall cylindrical shape containing a central cavity; the height and diameter of the cylinder are approximately 160 nm by 130 nm, respectively, and the subunits are arranged in two heptameric rings stacked on top of each other (Hendrix, 1979; Hohn *et al.*, 1979; Ishii *et al.*, 1992; Langer *et al.*, 1992b). Each

subunit contains two domains of approximately equal size which appear as a double-ring in electron microscope side views. The electron dense central channel, which traverses the protein, is approximately 6 nm in diameter and has been proposed to be the site for polypeptide binding (Langer *et al.*, 1992b; Braig *et al.*, 1993). GroEL also possesses a weak ATPase activity ($k_{cat} \approx 0.06$ sec^{-1} per subunit) (Hendrix, 1979) which is dependent on the presence of Mg^{2+} and K^{+} ions for maximal activity (Viitanen *et al.*, 1990). Nothing is known to date about the tertiary organisation of the GroEL polypeptide.

The *gro* E operon also encodes another heat-shock inducible protein known as GroES. GroES has a subunit M_r of 10, 368 (Hemmingsen *et al.*, 1988) and exists in a homo-oligomeric form composed of seven identical subunits arranged in a ring-like structure, 8 nm in diameter with a 2 nm electron dense central channel (Chandrasekhar *et al.*, 1986). GroES and its homologues in mitochondria (Hsp10) and plastids (cpn10), which are thought to have a similar structural arrangement, have also been defined as (co-)chaperonins (Hemmingsen *et al.*, 1988).

Genetic and biochemical evidence exists that the GroEL and GroES proteins functionally interact, which helps explain why mutations in either the *gro* ES gene or *gro* EL gene often results in a similar phenotype. The genetic evidence consists of the isolation of missense mutations in the *gro* EL gene which compensates for the temperature sensitive phenotype of certain *gro* ES mutations in an allele specific manner (Tilly and Georgopoulos, 1982). Biochemical evidence has shown that the uncoupled ATPase activity of GroEL is effectively inhibited by the binding of GroES to GroEL, an event that itself requires the presence of ATP or ADP (Chandrasekhar *et al.*, 1986; Viitanen *et al.*, 1990; Gray and Fersht, 1991; Bochkareva *et al.*, 1992). Such a stable GroEL-GroES complex can be isolated by sedimentation or gel filtration (Chandrasekhar *et al.*, 1986; Viitanen *et al.*, 1990; Bochkareva *et al.*, 1992) and is believed to be the basis of GroE function *in vivo*.

1.5.6.3 GroE homologues in higher organisms

The GroEL homologues identified in bacteria, fungi and plants all have the characteristic double-toroid tetradecameric structure observed in GroEL. The chloroplast homologue (ch-cpn60) is unusual, however, in that it is composed of two types of subunit of relative M_r 61, 000 and 60, 000, known respectively as the α and β subunits (Hemmingsen *et al.*, 1988). The two subunit types are highly divergent, with the α subunit showing a high sequence homology to GroEL, for example the wheat plastid α subunit of ch-cpn60 shows 46% identity to GroEL (Hemmingsen *et al.*, 1988). The two subunit types are present in equal amounts in the tetradecamer, but their structural arrangement within the chaperonin complex is unknown. A mammalian homologue of GroEL has recently been identified in bovine mitochondria (Viitanen *et al.*, 1992b). In contrast to previously characterised chaperonin 60 molecules this protein functions as a single heptameric toroid. The mammalian mitochondrial Hsp60 is virtually identical to GroEL in face views when examined by electron microscopy but the side view of GroEL is about twice as long, carrying the characteristic four striations running parallel to the long-axis compared to two such striations in bovine mitochondrial Hsp60 (Viitanen *et al.*, 1992b). It is unknown whether these mammalian chaperonin 60s generally function as a single-ring unlike the bacterial and plant chaperonin 60s which function as double rings.

Co-chaperonin, or *E. coli* GroES, homologues in higher organisms have been identified only relatively recently. These proteins, which are believed to interact functionally with their larger partner chaperonin proteins in a manner analogous to the GroE system, have been identified using a GroEL-dependent *Rhodospirillum rubrum* rubisco reconstitution assay. They are found to be present in mammalian mitochondria (Lubben *et al.*, 1990; Hartman et al., 1992), yeast mitochondria (Rospert *et al.*, 1993) and higher plant chloroplasts (Bertsch *et al.*, 1992). The stress-inducible rat liver

mitochondrial Hsp10 shares approximately 40% sequence homology with GroES and has a subunit M_r of 10, 813 (Hartman et al., 1992). The oligomeric arrangement of this protein, judging from gel filtration experiments, is likely to be hexameric or heptameric, like GroES (Hartman et al., 1992). The yeast mitochondrial Hsp10 protein has an approximate M_r of 9, 000, from SDS-PAGE, although nothing is known about its oligomeric arrangement (Rospert *et al.*, 1993). Unexpectedly this protein is unable to inhibit the ATPase activity of its partner chaperonin, yeast mitochondrial Hsp60, although it can form a stable, isolatable complex with it in the presence of ATP. This, therefore, may be a property unique to the GroE system. The GroES equivalent in higher plant chloroplasts, ch-cpn10, is apparently a 'double' chaperonin having a subunit size of approximately 24-kD which is twice the size of other reported chaperonin-10s (Bertsch *et al.*, 1992). This protein has apparently arisen through either gene duplication or through the fusion of two distinct chaperonin-10 genes. Both the N- and C-terminal halves of the mature protein exhibit high homology to GroES and the two halves show 40% identity (70% if conservative changes are taken into account) to each other (Bertsch *et al.*, 1992). The quaternary structure of this protein is not known with any certainty. Thus both the ch-cpn60 and ch-cpn-10 proteins are unique in their structural arrangement. These proteins may have evolved to perform specialised functions, perhaps related to the assembly of higher plant rubisco. As it is unknown whether multiple isoforms of ch-cpn60 exist one cannot rule out the possibility of a more conventional ch-cpn10 homologue.

1.5.6.4 Studies on GroE mediated (re)folding

The chaperonin proteins play an important role in stress protection under unbalanced physiological conditions, however it has become apparent that these proteins also have a more general role within the cell, namely in assisting the folding and / or assembly of a variety of proteins during the normal functioning of the cell The first evidence that chaperonin proteins were involved in a general role in protein folding

and / or assembly *in vivo* came from studies on a yeast strain (*mif*-4) which harboured a temperature sensitive mutation in the Hsp60 gene encoding the mitochondrial chaperonin Hsp60. Although protein synthesis and translocation into the mitochondrial matrix is unaffected in this strain at non-permissive temperatures polypeptides imported into the mitochondria, including Hsp60 itself, are unable to assemble into their native catalytically active structures (Cheng *et al.*, 1989; Cheng *et al.*, 1990). This is indicative of the importance of these proteins.

Chaperonins do not only interact with and promote the folding of polypeptides that have first crossed membranes, as studies on the GroE proteins have demonstrated. Overexpression of GroEL and its co-chaperonin GroES *in vivo* results in the enhanced expression of a number of foreign proteins in *E. coli* which would otherwise misfold or aggregate (Carrillo *et al.*, 1992; Lee and Olins, 1992; Wynn *et al.*, 1992; Edgerton *et al.*, 1993). This is consistent with the proposed role for the GroE proteins in protein biogenesis. Newly synthesised *E. coli* proteins are known to undergo aggregation in cells lacking a heat shock response (Gragerov *et al.*, 1992) but a more global role for the products of the *gro* E gene in polypeptide folding has recently been reported. Horwich *et al.* (1993), in line with the *mif*-4 mutants described above, introduced a temperature sensitive lethal mutation into the *gro* E gene and demonstrated that a defined group of cytoplasmic proteins, including citrate synthase, ketoglutarate dehydrogenase and polynucleotide phosphatase, are synthesised but fail to attain a native conformation when the cells are shifted to a non-permissive temperature. These authors estimated that approximately 30% of the soluble protein species are affected through GroEL deficiency; this is in rough agreement with the observations of Viitanen *et al.* (1992a) who estimate that approximately 50% of soluble protein species of *E. coli* are bound by GroEL upon dilution from chaotrope *in vitro*.

Much of the information regarding the mechanism of how GroEL functions in assisting protein folding and / or assembly have come mainly through *in vitro* studies

using purified chaperonins and chemically denatured purified 'substrate' proteins. GroEL has been shown to interact with a number of unrelated polypeptides *in vitro* and mediate the refolding into their final functional structure. Polypeptides bind to GroEL in an ATP independent process and generally form a stable complex with the chaperonin. Polypeptides can be released from the chaperonin in a folding-competent state *in vitro* in one of two ways. Release may be triggered by the presence of ATP alone or by the presence of GroES and ATP. In the former case GroEL has been shown to interact with and assist the refolding of a number of polypeptides including lactate dehydrogenase (Badcoe *et al.*, 1991; Jackson *et al.*, 1993), phytochrome photoreceptor (Grimm *et al.*, 1993), barnase (Gray and Fersht, 1993), enolase (Kubo *et al.*, 1993), glutamine synthase (Fisher, 1992), dihydrofolate reductase (Viitanen *et al.*, 1991; Martin *et al.*, 1992), tryptophanase (Mizobata *et al.*, 1992), the F_{ab} fragment of a monoclonal antibody (Schmidt and Buchner, 1992), α–glucosidase (Höll-Neugebauer *et al.*, 1991) and pre-β–lactaminase (Laminet *et al.*, 1990). It should be noted that whilst ATP is sufficient for the successful release and reactivation of these proteins, the presence of GroES potentiates the effect in most of these examples.

The proteins which have an essential requirement for both ATP and GroES for efficient GroEL mediated folding include rhodanese (Mendoza *et al.*, 1991; Mendoza *et al.*, 1992; Martin *et al.*, 1991; Bochkareva *et al.*, 1992; Tsalkova *et al.*, 1993), ornithine transcarbamylase (Zheng *et al.*, 1993) rubisco (Goloubinoff *et al.*, 1989; Viitanen *et al.*, 1990; Baneyx and Gatenby, 1992; van der Vies *et al.*, 1992; Viitanen *et al.*, 1992a), citrate synthase (Buchner *et al.*, 1991), malate dehydrogenase (Hartman *et al.*, 1993a; Miller *et al.*, 1993) and 6-hydroxy-D-nicotine oxidase (Brandsch *et al.*, 1992). The two best studied examples are rhodanese and rubisco. In each case the GroEL bound polypeptide can be released in the presence of ATP alone but fail to attain native conformation due to aggregation of the released polypeptides (Martin *et al.*,1991; Baneyx and Gatenby, 1992; Viitanen *et al.*, 1992a). In the presence of GroES and ATP, 85-90% of bound *Rhodospirillum rubrum* rubisco is released from GroEL

leading to the formation of both the active rubisco dimer and a smaller amount of inactive, but folded, monomer (Baneyx and Gatenby, 1992; Viitanen *et al.*, 1992a). In contrast, in the presence of ATP alone 50-75% of bound rubisco is released from GroEL but most of the material is unable to fold into either the monomeric or dimeric form (Baneyx and Gatenby, 1992; Viitanen *et al.*, 1992a). Similar results have been observed with the ATP induced release of rhodanese from GroEL, *i.e.* rhodanese released in the presence of ATP, but in the absence of GroES, is unable to attain native structure and undergoes aggregation (Martin *et al.*, 1991). The polypeptide is therefore released from GroEL in a different conformational state depending on whether ATP or ATP and GroES was responsible for release. In the presence of GroES, the bound polypeptide is able to progress to a state whereby it is not susceptible to aggregation on release and is committed to folding, indicating that the co-chaperonin plays a regulatory role in co-ordinating the timely release of polypeptides from GroEL.

1.5.6.5 GroEL as a catalyst

GroEL is undoubtedly a catatyst, albeit a sluggish one, with respect to the K^+ dependent hydrolysis of ATP to ADP and Pi. However the question of whether GroEL is a catalyst with respect to protein folding remains to be fully resolved. In most of the examples investigated the presence of the GroE proteins and ATP increases the yield of active protein upon refolding (often by suppressing aggregation). However in many of these cases little or no enhancement in the actual rate of (re)folding is observed (*e.g.* Laminet *et al.*, 1990; Buchner *et al.*, 1991; Badcoe *et al.*, 1991; Höll-Neugebauer *et al.*, 1991; Schmidt and Buchner, 1992; Gray and Fersht, 1993; Kubo *et al.*, 1993; Miller *et al.*, 1993; Jackson *et al.*, 1993). In the case of dihydrofolate reductase the rate of refolding in the presence of GroEL and ATP is actually slower than that observed for refolding in the absence of chaperonins. The presence of GroES slightly increases the refolding rate, although it is still slower than spontaneous refolding (Martin *et al.*, 1991; Viitanen *et al.*, 1991). In contrast only two cases have been reported to date

whereby the rate of refolding is accelerated in the presence of the chaperonin system, namely the refolding of rubisco (Viitanen *et al.*,1990) and ornithine transcarbamylase (Zheng *et al.*, 1993). The rate of spontaneous rubisco refolding at 15°C is increased approximately 10 fold in the presence of the GroE proteins and ATP (Viitanen *et al.*,1990). It is difficult to interpret the results of such experiments with respect to what occurs *in vivo* , since the success of rubisco refolding *in vitro* (and that of other proteins) depends on the experimental conditions chosen, regarding ionic strength and temperature. The majority of the evidence available to date would, however, suggest that GroEL does not act as a catalyst of protein refolding (*i.e.* to enhance the rate of refolding) but rather acts to prevent off-pathway interactions which may inhibit correct folding.

1.5.6.6 Recognition of target polypeptides by GroEL

GroEL interacts with a variety of polypeptides which are unrelated with regards to sequence, structure and function and assists in their (re)folding to native conformation (see section 1.5.6.4). With two exceptions, pre-β–lactaminase (Laminet *et al.*, 1990) and dihydrofolate reductase (Viitanen *et al.*, 1991), GroEL does not interact with polypeptides once they have attained native structure but rather only interacts with their non-native forms. Thus the structural motif(s) recognised by GroEL must be present only during the early stages of protein (re)folding and be no longer present or inaccessible in the native protein. Indeed in the case of pre-β–lactaminase (Laminet *et al.*, 1990) and dihydrofolate reductase (Viitanen *et al.*, 1991) these proteins are known to exist in a slow equilibrium with a number of conformational states, at least one of which may interact with GroEL.

Studies by Martin *et al.* (1991) have indicated that both rhodanese and dihydrofolate reductase are bound to GroEL in a similar conformation. These proteins, when bound to GroEL, are more susceptible to proteolysis (compared to the native

protein), show enhanced binding of the hydrophobic probe ANS, and show fluorescence properties intermediate between that of the fully unfolded state and the native state. Thus GroEL stabilises rhodanese and dihydrofolate reductase in a conformation lacking the ordered tertiary structure characteristic of an early folding intermediate, the 'molten globule' (see section 1.3). Indeed some partial folding may also occur whilst the polypeptide is associated with GroEL (Martin *et al.*, 1991). Although Badcoe *et al.* (1991) have proposed that GroEL does not interact with the 'molten globule' form of lactate dehydrogenase, but rather an earlier folding intermediate, many other polypeptides are believed to interact with GroEL in such a conformation. Independent studies on GroEL-bound rhodanese (Mendoza *et al.*, 1992) have confirmed the proposal of Martin *et al.* (1991) whilst the fluorescence spectra of GroEL-bound α–glucosidase is suggestive of a 'molten globule' type structure (Höll-Neugebauer *et al.*, 1991). Evidence also exists that GroEL interacts with a compact intermediate form of rubisco (van der Vies *et al.*, 1992) whilst an intermediate of the α–subunit of luciferase, which lacks stable tertiary structure (as judged by N.M.R.) but contains a high degree of secondary structure, is recognised by GroEL (Flynn *et al.*, 1993).

The 'molten globule' folding intermediate contains a high proportion of exposed hydrophobic residues compared to the native protein and as such may have a propensity to aggregate in the absence of chaperonins. Thus it is conceivable that GroEL may, in part, recognise and target hydrophobic residues as its site of interaction. The answer to this question will come from high resolution techniques such as N.M.R. studies but to date only limited such studies have been undertaken. A peptide corresponding to the N-terminal 13 amino acids of rhodanese binds to GroEL in an α–helical conformation (Landry and Gierasch, 1991). These authors have proposed that GroEL interacts with peptides that have the potential to adopt amphipathic α–helixes, perhaps by interacting with a hydrophobic face. Interestingly another synthetic peptide, vsv-C, binds to GroEL in a α–helical form but binds to DnaK (the *E. coli* Hsp70 homologue) in a

more extended form (Landry *et al.*, 1992). This is likely to reflect the fact that chaperonins are involved further downstream in the folding pathway than the Hsp70s. However, GroEL does not just bind to (poly)peptides which have the potential to adopt α–helixes. Schmidt and Buchner (1992) have shown that GroEL can functionally interact with an F_{ab} fragment of a monoclonal antibody which is an all β–protein. Therefore, as yet, no 'typical' consensus sequence has been observed for GroEL binding as has been observed for members of the Hsp70 class (Blond-Elgundi *et al.*, 1993a; section 1.5.2.2.). It is likely that polypeptide interactions with GroEL will depend mainly on the nature of the folding intermediate rather than a specific element of secondary structure.

1.5.6.7 The stoichiometry of binding and the site of recognition

The stoichiometry of binding of unfolded polypeptides to GroEL has been examined mainly by measuring the amount of GroEL required to prevent aggregation or spontaneous refolding of polypeptides on dilution from denaturant (Laminet *et al.*, 1990; Badcoe *et al.*, 1991; Martin *et al.*, 1991; Mendoza *et al.*, 1991; Bochkareva *et al.*, 1992). The results of such studies have indicated that one or at most two molecules can be bound per GroEL tetradecamer. Direct examination of gold-labelled dihydrofolate reductase molecules bound to GroEL suggest that two molecules (but occasionally three) bind to GroEL (Braig *et al.*, 1993).

The stoichiometry of one or two molecules binding per GroEL oligomer may suggest that binding occurs at the centre of the GroEL complex. Indeed scanning transmission electron microscopy studies have demonstrated that unfolded dihydrofolate reductase tagged with gold clusters binds within the central cavity of GroEL (Braig *et al.*, 1993). In support of this, negative stain electron microscopy images of both rhodanese and alcohol oxidase bound-GroEL show a stain excluding mass in the central portion of GroEL which is not present in the absence of 'substrate'

polypeptides (Langer *et al.*, 1992b). Thus it seems as though 'substrate' polypeptides do indeed bind within the central cavity of GroEL. This cavity can accomodate polypeptides up to approximately 90-kD (Braig *et al.*, 1993) and conveys several advantages, namely (a) folding intermediates can bind at multiple sites of the surrounding monomeric subunits and upon GroES / ATP mediated release be rebound by the surrounding monomers (or by the adjacent ring) if a native-like conformation is not initially achieved, (b) by physically removing the polypeptide from the complex cellular environment its chances of aggregating is greatly reduced and (c) with the polypeptide binding sites of GroEL positioned within a cavity, homotypic interactions between GroEL rings themselves would be minimised. The location of the polypeptide along the long-axis of GroEL is not fully understood, however it is interesting to note that a functional Hsp60 chaperonin has been isolated from mammalian mitochondria which, unlike other characterised chaperonin-60 molecules, exists as a single heptameric toroid (Viitanen *et al.*, 1992b) (see section 1.5.6.3 above).

1.5.6.8 GroEL - GroES interactions

In vitro studies have indicated that whilst some proteins can refold efficiently in the absence of GroES, for others the presence of GroES is absolutely essential to attain native conformation (see section 1.5.6.4). This may, in part, reflect the relative binding affinities of GroEL for different polypeptides and / or indicate that slower folding polypeptide intermediates require additional interactions with GroEL, in a GroES and ATP-assisted process, before being released in a folding-competent state. It should also be noted that GroEL does not assist in the refolding of all polypeptides within the cell (Horwich *et al.*, 1993). Such *in vitro* experiments, however, do not rule out a role for GroES *in vivo* and in many cases where GroES was not essential for efficient release and refolding it did potentiate the effect (Laminet *et al.*, 1990; Höll-Neugebauer *et al.*, 1991; Viitanen *et al.*, 1991; Fisher, 1992; Martin *et al.*, 1992; Schmidt and Buchner,

1992; Gray and Fersht, 1993). Thus *in vivo* GroEL and GroES are likely to function together.

The stoichiometry of (ATP / ADP dependent) GroES binding to GroEL is one GroES heptamer per GroEL tetradecamer. Several lines of evidence have determined this, for example maximum inhibition of the uncoupled GroEL ATPase activity by GroES occurs at a 1 : 1 stoichiometry (Chandrasekhar *et al.*, 1986) whilst the *in vitro* refolding of polypeptides occur optimally when the chaperonins are present together in such a ratio (Goloubinoff *et al.*, 1989; Martin *et al.*, 1991; Langer *et al.*, 1992b; Miller *et al.*, 1993). Electron microscopy image analysis has directly demonstrated such GroEL - GroES interactions both *in vitro* (Saibil *et al.*, 1991; Langer *et al.*, 1992b) and *in vivo* (Ishii *et al.*, 1992). A single GroES ring binds asymmetrically to one end of the GroEL cylinder producing a 'bullet-like' appearance (Saibil *et al.*, 1991; Ishii *et al.*, 1992; Langer *et al.*, 1992b). The binding of the co-chaperonin to GroEL induces a conformational change in both the GroES adjacent end and the opposite end of the GroEL cylinder (Langer *et al.*, 1992b). This apparently prohibits the binding of a second GroES oligomer. Interestingly, GroES can interact initially with either end of the GroEL cylinder (Langer *et al.*, 1992b), which may suggest that the GroEL rings are arranged in an A-B-B-A formation along the long-axis (compared to an A-B-A-B formation). The nature of the GroEL-GroES interaction remains unclear. Landry *et al.* (1993) have postulated that native GroES contains a highly flexible and accessible polypeptide loop, spanning amino acid residues 17-32 in the GroES sequence, which is important for binding to GroEL at a GroES specific site. This site be may located near the C-terminus of GroEL since GroES binding prevents proteolytic cleavage of the C-terminal 16 residues of GroEL (Langer *et al.*, 1992b). Further characterisation will be required, however, to identify the molecular details of such an interaction.

GroEL can also functionally interact with the GroES homologues identified to date from higher organisms. GroEL can form stable, isolatable complexes in the

presence of ATP with bovine mitochondrial Hsp10 (Lubben *et al.*, 1990), cpn10 from pea chloroplast (Bertsch *et al.*, 1992) and yeast mitochondrial Hsp10 (Rospert *et al.*, 1993). In each case, these heterologous chaperonin complexes are able to assist in the refolding of denatured *Rhodospirillum rubrum* rubisco in a manner analogous to the GroE complex. Refolding of ornithine transcarbamylase by a heterologous GroEL / rat mitochondrial Hsp10 system has also been reported (Hartman *et al.*, 1992). This is likely to reflect a considerable conservation of interactive surfaces on GroEL and the diverse chaperonin-10s. Interestingly, although the cpn10 of higher plant chloroplasts is partially functional with yeast mitochondrial Hsp60 (Rospert *et al.*, 1993), GroES can functionally interact with neither yeast mitochondrial Hsp60 (Rospert *et al.*, 1993) nor with bovine mitochondrial Hsp60 (Viitanen *et al.*, 1992b), either to form a stable heterologous chaperonin complex or to promote rubisco refolding. No interactions between GroES or plastid cpn60 has been reported either. The complete primary sequence of rat mitochondrial chaperonin-10 (Hsp10) has recently revealed that, unlike GroES, the amino terminal of the polypeptide is acetylated (Hartman *et al.*, 1993b). It is feasible, therefore, that N^{α}–acetylation of rat Hsp10 (and therefore of other chaperonin-10s) is required for a functional interaction with its mammalian Hsp60 counterpart.

In summary it is likely that GroEL and GroES functionally interact *in vivo*, certainly under normal cellular conditions. The role of the GroES protein is to co-ordinate the ATP dependent release of bound segments of the 'substrate' polypeptide and prevent premature release which may lead to aggregation. Under heat -shock conditions, *i.e.* at high temperatures, some GroEL can become reversibly phosphorylated which allows ATP-dependent release of substrate polypeptides to occur in the absence of GroES (Sherman and Goldberg, 1992). This form of GroEL may be more efficient, than the major GroEL species, in promoting the repair of (heat) damaged polypeptides

1.5.6.9 The ATP binding site of GroEL

The GroEL oligomer contains 14 ATP-binding sites and binding and hydrolysis of the nucleotide is a highly co-operative process which is enhanced by the presence of GroES (Gray and Fersht, 1991; Bochkareva *et al.*, 1992; Jackson *et al.*, 1993). GroES is believed to act to couple the hydrolysis of ATP with the release of the polypeptide in a folding competent state (Viitanen.*et al.*, 1990). The molecular mechanism of this tightly coupled interaction is not yet fully understood. On a simplistic level polypeptides, at least *in vitro*, can be released from GroEL simply through the binding of ATP. ATP hydrolysis is not required to alter the conformational structure of GroEL to reduce its affinity for 'substrate' proteins. GroEL undergoes similar conformational change in the presence of ATP, ADP or non-hydrolysable ATP analogues (Baneyx and Gatenby, 1992) and many polypeptides which interact with GroEL can be released in the presence of non-hydrolysable ATP analogues or ADP (table 1.2) and fold into their native state. The refolding yield is generally lower than in the presence of ATP and reflects the fact that the ATP 'substitutes' bind more weakly to the ATP-binding site (Baneyx and Gatenby, 1992; Jackson *et al.*, 1993). Interestingly almost all of these proteins do not have essential requirements for GroES to assist protein refolding *in vitro*. The exception is ornithine transcarbamylase which normally requires GroES and ATP for efficient refolding. This protein can be released from GroEL in the presence of ATPγS and GroES but the released polypeptide is not competent for assembly (Zheng *et al.*, 1993). It appears that in those cases where an absolute requirement for GroES exists, there is also an absolute requirement for hydrolysis of ATP. Indeed the energy requirement of GroEL-GroES assisted rhodanese refolding *in vitro* is relatively high. About 130 molecules of ATP are hydrolysed per molecule of rhodanese folded, indicating that the reaction cycle undergoes multiple rounds of ATP hydrolysis per GroEL tetradecamer (Martin *et al.*, 1991). During such a folding process the polypeptide substrate and the regulatory protein GroES undergo cycles of binding and

Table 1.2

Polypeptides which can be released from GroEL in the absence of ATP hydrolysis

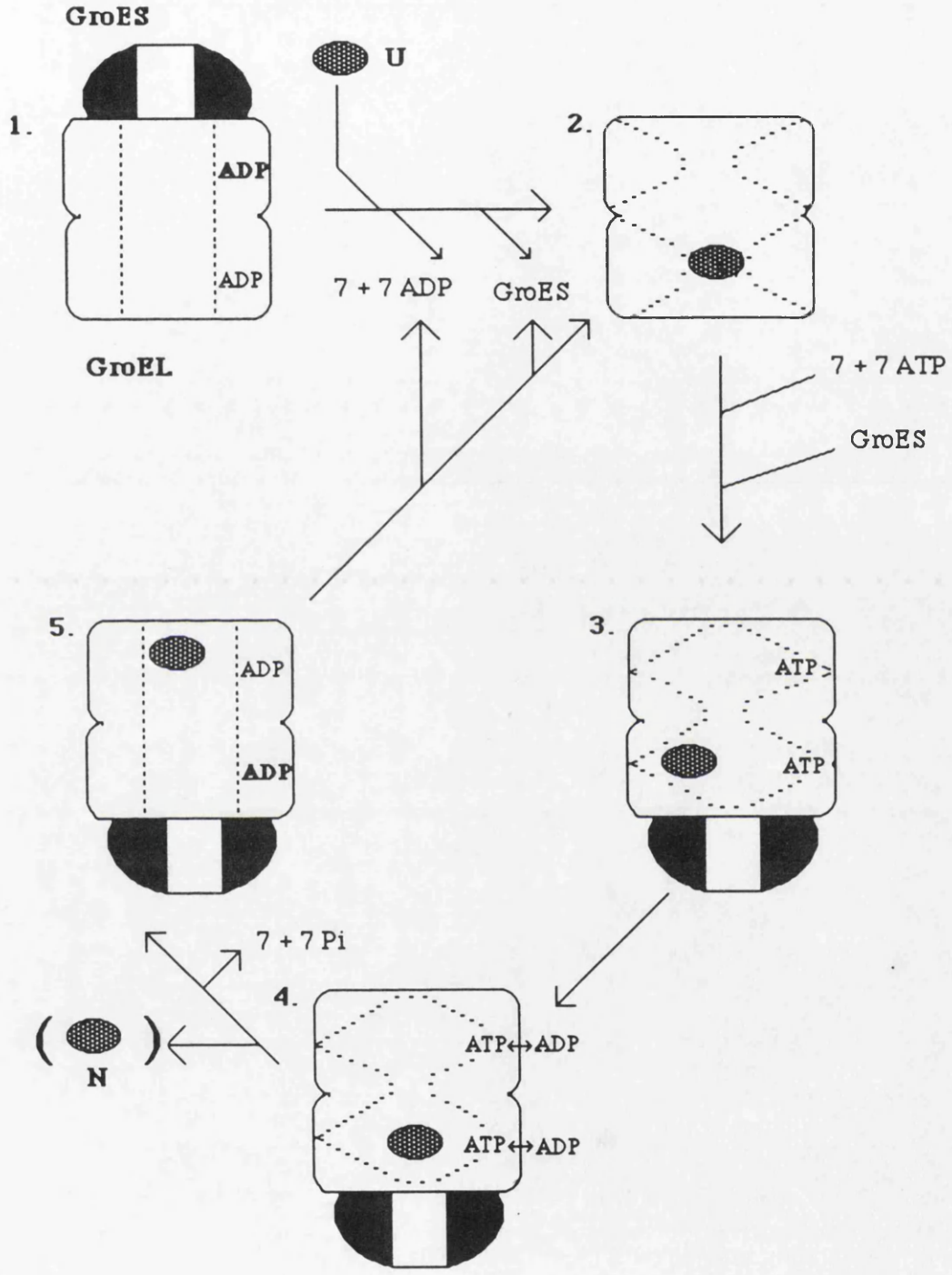
Protein	Releasing agent(s)	Reference
barnase	AMP-PNP, ADP	Gray and Fersht (1993)
dihydrofolate reductase	AMP-PNP, ATPγS	Viitanen *et al.* (1991)
enolase	ADP, CTP, UTP	Kubo *et al* (1993)
F_{ab} fragment	AMP-PNP, ATPγS	Schmidt & Buchner (1991
glutamine synthase	AMP-PNP, ATPγS	Fisher *et al.* (1992)
lactate dehydrogenase	AMP-PNP	Badcoe *et al.* (1991)
ornithine transcarbamylase	AMP-PNP	Zheng *et al.*(1993)
phytochrome photoreceptor	ATPγS	Grimm *et al* (1993)
tryptophanase	AMP-PNP, ATPγS, ADP	Mizobata *et al.* (1992)

release to GroEL until polypeptide folding is complete (Martin *et al.*, 1991; Martin *et al.*, 1993a).

A general model for the chaperonin-mediated folding pathway can be proposed (see Fig. 1.1). This model is based on the existence of both 'high' and 'low' affinity states of GroEL for either polypeptide binding or GroES binding which are mutually exclusive. In the absence of substrate polypeptides the ATPase activity of GroEL is effectively inhibited by the binding of GroES (Chandrasekhar *et al.*, 1986; Viitanen *et al.*, 1990; Gray and Fersht, 1991; Bochkareva *et al.*, 1992), which presumably prevents wasteful energy release. This complex is probably the preferred state under physiological conditions and GroEL is stabilised by GroES in a form with a high affinity for ADP (Jackson *et al.*, 1993; Martin *et al.*, 1993a). Addition of unfolded polypeptide to such a complex is known to result in a stimulation of the GroEL ATPase activity (Martin *et al.*, 1991; Jackson *et al.*, 1993). The 'substrate' polypeptide presumably binds to the free cavity opposite the bound GroES molecule. The binding of the unfolded polypeptide into the central cavity of GroEL reduces its affinity for ADP and as a consequence both ADP and the co-chaperonin are released (Martin *et al.*, 1993a). This in turn makes the seven subunits in the interacting GroEL toroid accessible for ATP binding. The binding of ATP lowers the affinity of GroEL for bound polypeptides and weakly bound substrate proteins may be released (either completely or partially) at this point. Tightly bound polypeptides, however, require the (GroES-assisted) hydrolysis of ATP for productive release. GroES reassociates with the ATP-bound GroEL complex, binding to the opposite ring, prior to ATP hydrolysis (Martin *et al.*, 1993a). GroEL then undergoes rapidly consecutive rounds of ATP hydrolysis, in both rings, causing release of the polypeptide which is then free to undergo folding in the central cavity. Proteins which have folded or have attained a folding-competent state during this single cycle are free for release. This also applies to the folded subunits of oligomeric proteins before their assembly. Proteins which have

Fig. 1.1 Model for the GroEL-GroES assisted folding of polypeptides

See text for main details and references. (1) The GroEL-GroES complex is favoured under physiological conditions. The GroES associated ring of GroEL is stabilised in a high affinity ADP-binding state (denoted **ADP** in bold), the other ring has a lower affinity for ADP. (2) Upon binding of the unfolded polypeptide (**U**), ADP and consequently GroES dissociate. (3) ATP binds, weakening the interaction between the polypeptide and GroEL, and allows GroES to rebind. (Proteins with a low affinity for GroEL may be released at this stage). (4) ATP hydrolysis releases the bound polypeptide into central cavity, allowing it to fold. Proteins attaining (near) native structure after this single cycle are released (*i.e.* **N**). (5) Generation of the ADP state increases the binding affinity of GroEL for GroES. Polypeptides which are still (partially) unfolded rebind and the reaction cycle continues (from step 2) until a folding competent state is reached. Between steps (5) and (2) the GroEL protein is rotated 180°C, with the substrate polypeptide remaining bound in the same ring. This model was adapted from Martin *et al.* (1993a).


GroES
1.
ADP
ADP
GroEL
U
7 + 7 ADP
GroES
2.
7 + 7 ATP
GroES
5.
ADP
ADP
3.
ATP
ATP
7 + 7 Pi
N
4.
ATP↔ADP
ATP↔ADP

only partially folded and / or have exposed recognition elements may then rebind to the ADP-stabilised GroEL - GroES complex and undergo further cycles of release and folding until a (near) native conformation is released. It is unknown whether the folding polypeptide can shuttle between the two rings of the chaperonin during folding and it remains for the molecular details of this reaction cycle to be characterised further.

1.6 Sequential action of the chaperone Hsp70 and Hsp60 proteins

The discussion of the various classes of chaperone proteins separately may give the misleading impression that such proteins play specialised but independent roles within the cell. However, the Hsp70 and Hsp60 chaperones have distinct but complementary roles with respect to polypeptide folding. The Hsp70s act as monomers or dimers within the cell and their main role during protein folding seems to be in shielding exposed hydrophobic surfaces of nascent polypeptides during translation (or during translocation) until at least a complete protein domain is available for productive folding (section 1.5.2). The Hsp60s, on the other hand, act as large oligomeric double-ring complexes which sequester completely synthesised but unfolded polypeptides from the complex cellular milieu and not only prevent aggregation but help promote folding into the native conformation (section 1.5.6). These distinct classes of chaperone proteins act sequentially in the folding pathway. Polypeptides entering the mitochondrial matrix, in a loosely folded conformation resembling the nascent chain emerging from the ribosome, interact initially with matrix-located Hsp70 proteins (Kang *et al.*, 1990; Scherer *et al.*, 1990) prior to transfer to mitochondrial Hsp60 which, at least in some cases, is responsible for subsequent folding to attain native structure (Cheng *et al.*, 1989; Ostermann *et al.*, 1989; Cheng *et al.*, 1990). The basis for this differential recognition is the distinct structural features recognised by both the Hsp70 and Hsp60 proteins. The Hsp70 proteins recognise polypeptides in an extended form which have little or no secondary structure (section 1.5.2.2), whilst members of

the Hsp60 class recognise early folding intermediates of polypeptides resembling 'molten globules' (section 1.5.6.6).

The transfer of polypeptides from Hsp70 to Hsp60 proteins is not a feature unique to mitochondria. Indeed transfer has been demonstrated *in vitro* between the respective *E. coli* cytoplasmic homologues (Langer *et al.*, 1992a). The *E. coli* Hsp70 homologue DnaK in conjunction with its partner chaperone DnaJ stabilises rhodanese in an unfolded conformation and transfer to GroEL, for subsequent folding, is mediated by the 'coupling-factor' GrpE and ATP (Langer *et al.*, 1992a). It remains to be established whether this pathway occurs *in vivo* since both DnaK and DnaJ, unlike the GroE proteins, are dispensable for *E. coli* at normal growth temperatures (Bukau and Walker, 1989; Sell *et al.*, 1991), although it is possible that the GroE system can function alone or that a second pair of DnaK / DnaJ related proteins exist which may be able to function generally, albeit with a low(er) efficiency. Recent evidence has, however, suggested that DnaK and GroEL have complementary effects in preventing polypeptide aggregation within the cell (Gragerov *et al.*, 1992). It should also be noted that not all polypeptides fold by such a sequential chaperone pathway, consistent with the observation that some *E. coli* cytosolic proteins can fold in the absence of GroEL (Horwich *et al.*, 1993). Some polypeptides may be free to fold and attain a native conformation simply on release from the Hsp70 chaperone. Indeed reactivation of thermally denatured firefly luciferase requires a functional DnaK system (*i.e.* DnaK / DnaJ / GrpE / ATP) but not the GroE chaperonins (Schröder *et al.*, 1993). This sequential model for the folding pathway of (some) polypeptides remains to be confirmed *in vivo* and demonstrated in cell-free translation systems *in vitro*. Its relevance to polypeptide folding in the eukaryotic cytosol remains unclear as, to date, no GrpE homologues or potential co-operating factors for the chaperonin 60-like TRiC family of proteins have been identified. It remains a task for future work to characterise this interesting pathway further.

1.7 Aims of the project

The main aims of the PhD project were to study two unanswered but fundamental questions about the function of the GroEL protein, and were as follows;

(1) To identify the important amino acid residue(s) involved in the ATP binding site, in order to gain some information towards understanding the tertiary arrangement of the chaperone protein.

(2) To study the unfolding and the refolding of GroEL in order to gain some insight into how the GroEL protein, itself, attains native structure.

Chapter 2 Materials and Methods

2.1 Materials and reagents

2.1.1 Chemicals and biochemicals

Ampicillin, 8-anilino-1 naphthalene sulphonate (ANS), benzamidine, bromophenol blue, Coomassie brilliant blue, 5, 5'-dithiobis (2-nitrobenzoate), 5' p fluorosulphonyl benzoyladenosine, 4-(iodoacetamido)salicylic acid, sodium borohydride ($NaBH_4$), N,N,N',N'-tetramethylethylene diamine (TEMED), triethanolamine and triton X-100 were obtained from Sigma Chemical Co., Poole, Dorset, U.K.

Bactotryptone, yeast extract and Bactoagar (agar) were obtained from Difco, Detroit, Michigan, U.S.A.

ATP, DTT, NADH, phenylmethanesulphonyl fluoride (PMSF) and phospho*enol* pyruvate were obtained from Boehringer Mannheim, Lewes, Sussex, U.K.

2-Mercaptoethanol and sodium periodate were obtained from BDH Chemicals, Poole, Dorset, U.K.

Guanidinium chloride (GdnHCl) of ultrapure grade was obtained from BRL, Gibco Ltd., Paisley, Scotland, U.K.

Acrylamide, N, N' methylenebisacrylamide, SDS and trichloroacetic acid were obtained from FSA Laboratory Supplies, Loughborough, Leicestershire,U.K.

2',3'-dialdehyde ATP (oATP) was prepared by periodate oxidation of ATP following the procedure of Easterbrook-Smith *et al.* (1976). Purity was judged by thin

* Thin layer chromatography was performed on polyethyleneimine sheets using 0.8M-NH_4HCO_3 as the developing solution and u.v. light to detect the position of the nucleotide. The relative R_f values were 0.52 and 0.05 for ATP and oATP, respectively.

layer chromatography, as described, and the resulting product was divided into aliquots and stored at - 70 °C.

HPLC grade trifluoroacetic acid (TFA) and acetonitrile were obtained from Rathburn Chemicals Ltd., Walkerburn, Scotland, U.K.

[2',5',8-^{3}H] ATP was obtained from Amersham International Plc., Amersham, Buckinghamshire, U.K.

Polyvinyl difluoride (PVDF) membranes (Immobilin-P transfer membranes, as described by Matsudaira, 1987) were obtained from Millipore, Bedford, Massachusetts, U.S.A.

All other chemicals were of analytical grade and were obtained from one of the following suppliers: Aldrich Chemical Co. Ltd., Poole, Dorset, U.K.; BDH Chemicals; FSA Laboratory Supplies; Koch-Light Ltd., Colnbrook, Buckinghamshire, U.K.

2.1.2 Enzymes and proteins

Bovine serum albumin (BSA), pyruvate kinase / lactate dehydrogenase enzyme suspension, thermolysin (protease type X) and a molecular weight standard marker kit were obtained from Sigma Chemical Co.

Proteinase K was obtained from Boehringer Mannheim.

Trypsin was obtained from Koch-Light Ltd.

Chymotrypsin was obtained from Worthington Biochemical Corp., Freehold, New Jersey, U.S.A.

2.1.3 Chromatography media

Sephacryl S-200, Sephadex G-10 and Sephadex G-25 were obtained from Pharmacia, Milton Keynes, Buckinghamshire, U.K.. DEAE-cellulose (DE52) was obtained from Whatman Biochemicals, Maidstone, Kent, U.K.

2.1.4 Pre-packed media

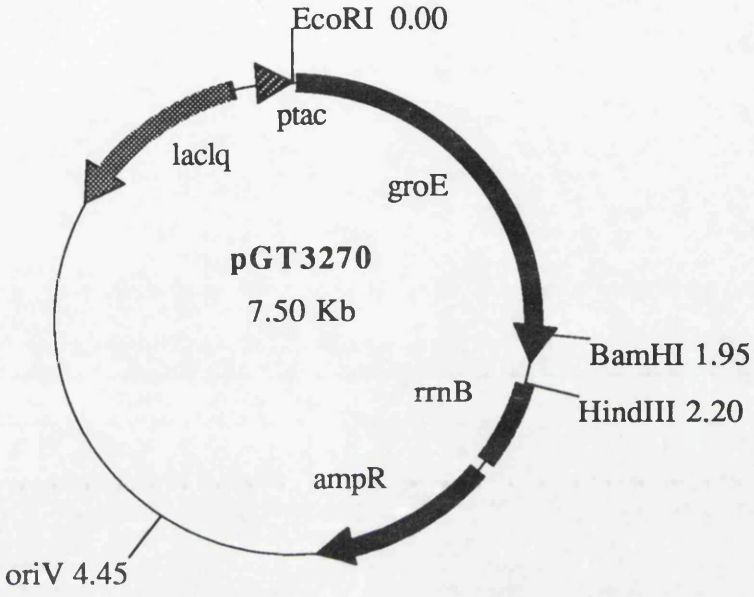
Pre-packed Mono Q, alkyl-superose and superose 6 columns were obtained from Pharmacia and utilised on a Pharmacia FPLC System. μBondapak C18 reverse phase column was obtained from Waters Chromatography, Watford, Hertfordshire, U.K.and attached to a Beckman System Gold HPLC apparatus (Beckman Instruments Inc., High Wycombe, Buckinghamshire, U.K.).

2.1.5 Bacterial strains

The bacterial strain used during this study was *E.coli* DH1 carrying either the plasmid

(a) pND5, an 8.1kb *Eco* R1 *E.coli* chromosomal fragment carrying the *gro* ESL operon (Jenkins *et al.*, 1986), or (b) pGT3270, containing the *gro* E genes on a 2.1kb *Eco* R1-*Hind* III DNA fragment (McLennan *et al.*, 1993) (Fig. 2.1). These strains were gifted by Dr. N.F. McLennan, ICMB, University of Edinburgh, Scotland, U.K.

Bacterial stocks were stored at -70°C as a frozen stock, prepared by mixing 5ml of overnight L-Broth (LB) culture with an equal amount of sterile glycerol. Short term storage was on a suitable sealed L-agar plate at 4°C.



Plasmid name: pGT3270
Plasmid size: 7.50 kb
Constructed by: NF McLennan
Construction date: unknown
Comments/References: Fig. 2.1 Plasmid pGT3270

2.2 General laboratory methods

2.2.1 pH measurement

pH measurements were carried out using a Corning 220 pH meter (Ciba Corning Diagnostics Ltd., Sudbury, Suffolk, U.K.), calibrated at room temperature.

2.2.2 Protein estimation

Protein concentrations were determined by the method of Bradford (1976), with BSA as standard.

2.2.3 Lyophilization

Protein and peptide solutions were collected either into Reacti Vials (Pierce Chemical Co., Rockford, Illinois, U.S.A.), polypropylene tubes (Eppendorf) or 25-100ml acid washed round-bottom flasks and the contents frozen by immersing the vessels in a dry-ice/ethanol mixture before lyophilization on a FTS Systems (Stone Ridge, New York, U.S.A.) Flexi-Dry freeze dryer.

2.3 SDS-polyacrylamide gel electrophoresis

Electrophoresis in the presence of SDS was performed by the method of Laemmli (1970), with a 3% stacking gel and a 12% or 15% separating gel. The ratio of acrylamide : bisacrylamide in all PAGE experiments was 30 : 0.8 and polymerisation was induced by the addition of 0.03% (v/v) TEMED and 0.05% (w/v) ammonium persulphate. After electrophoresis gels were stained for protein by the Coomassie method (Section 2.3.1)

PhastGels (Pharmacia) and SDS buffer strips, utilizing the Pharmacia Phast-System, were routinely employed for SDS-PAGE during GroEL purification.

2.3.1 Staining for protein

After electrophoresis, protein on the gels was visualised by staining with Coomassie blue for at least 1 hour at 40°C. The Coomassie reagent was 0.1% (w/w) Coomassie brilliant blue G250 in 50% (v/v) methanol, 10% (v/v) glacial acetic acid; destaining was carried out at 40°C in 10% (v/v) methanol, 10% glacial acetic acid until the background was fully destained and bands clearly visible.

2.4 Enzyme assays

Enzyme assays were performed at 20°C for unfolding / refolding experiments and at 25°C for chemical modification , in a total volume of 1ml, using a Gilford/Unicam model 252 uv/vis spectrophotometer. All ATPase assays were performed at least twice. Results were found to be highly reproducible between separate experiments.

2.4.1 GroEL ATPase assays

(a) GroEL ATPase coupled continuous assay

The ATPase activity of GroEL was measured spectrophotometrically by coupling the hydrolysis of ATP to the oxidation of NADH by the coupling enzymes pyruvate kinase and lactate dehydrogenase. Unless otherwise stated assays were performed in 50 mM-triethanolamine buffer pH 8.0 (with KOH) containing 10 mM-magnesium acetate. Concentrations of substrates and coupling enzymes in the buffer described were; 0.5 mM-ATP, 1.0 mM-phospho*enol* pyruvate, 0.15 mM-NADH, 25 µg/ml-pyruvate kinase and 8.3 µg/ml-lactate dehydrogenase. The oxidation of NADH was monitored at 340 nm (E_{340nm} = 6.22 x 10^3 M^{-1} cm^{-1}) Price *et al.* (1993).

(b) GroEL ATPase "coupled quench" assay

The effect of GdnHCl on the ATPase activity was studied by a coupled quench assay procedure of the type described by Johnson and Price (1988). 0.2 ml aliquots of the reaction mixture, containing GroEL in the presence of increasing concentrations of GdnHCl (Section 2.8.2) and 250μM-ATP, were added to 0.8 ml samples of the assay cocktail described above (Section 2.4.1 (a)). The rapid decline in A_{340} corresponded to the concentration of ADP formed during the ATPase reaction. By taking aliquots at 20 min intervals, the rate of the ATPase reaction was calculated. Control experiments of the type described by Johnson and Price (1988) were undertaken to confirm the validity of the method.

2.5 Growth of *E.coli* DH1/pGT3270 and *E.coli* DH1/pND5

2.5.1 Growth media

The 'rich' media used for *E.coli* growth in these studies was:

Medium	Composition (per litre)
L-broth (LB)	10 g bactotryptone 5 g yeast extract 10 g NaCl (+ 5 ml 20% (w/v) glucose)
L-agar	as LB, +15 g agar

All media was sterilised by autoclaving at 15 psi.

2.5.2 Selection supplements

Ampicillin (Amp) was added to the 'rich' media at a final concentration of 50 μg/ml, to select and identify the recombinant organisms. A stock solution of 25 mg/ml was filter sterilised and stored at -20°C. Hot L-agar was cooled to 55°C before ampicillin was added. L-Amp plates were stable for at least 4 weeks when stored at 4°C.

2.5.3 Cell growth

All cell growth was performed at 37°C, with continuous shaking.

E.coli DH1/pGT3270 or *E.coli* DH1/pND5 from a glycerol stock, was streaked out on an L-agar plate containing 50 μg/ml ampicillin and incubated overnight. 10 ml LB/amp was then inoculated with a single colony of *E.coli* and grown overnight. 2.5ml of this culture was then added to each of two 150 ml culture flasks of LB/amp. These cultures were grown overnight and 25 ml used to inoculate each of 8 x 500 ml LB/amp media, in 2 l conical flasks. Incubation continued until the cell suspension had reached saturation (OD_{600} of approximately 5) before harvesting the cells by centrifugation (4,500 g for 15 minutes using a M.S.E. Mistral 6L centrifuge). Cells which were not used immediately were stored at -20°C. A typical 4 l culture gave 24 g wet weight of cells.

2.5.4 Cell breakage

Cell pellets were resuspended in 2 volumes of ice-cold extraction buffer (50 mM-trisHCl pH 8.0, 5 mM-2-mercaptoethanol, 1 mM-PMSF, 1 mM-benzamidine) and broken by 2 passages through an automatic French pressure cell at 98 MPa (14 300 psi

internal pressure). The cell was pre-cooled on ice before use (cat. no. 4-3398A, American Instruments Company, Silver Spring, Maryland, U.S.A.).

2.6 GroEL purification

GroEL was purified from either *E.coli* DH1/pGT3270 or *E.coli* DH1/pND5 using a procedure modified from that described previously (Hendrix,1979; Chandrasekhar *et al.*, 1986). The protocol is fully detailed in section 3.2.

2.6.1 Storage of purified GroEL

Purified GroEL was dialysed for 24 hours against 100 volumes of 50 mM-trisHCl pH7.5 containing 50% (v/v) glycerol and stored at -20°C.

2.7 Determination of GroEL concentration

The characterisation and quantification of GroEL was performed spectrophotometrically using a value of 0.285 for the specific absorption coefficient (litre.g^{-1}.cm^{-1}) at 280 nm (Price *et al.*, 1991).

2.8 Unfolding and refolding of GroEL

2.8.1 Concentration of Guanidinium chloride

The concentrations of GdnHCl solutions were checked by refractive index measurements in accordance with the data given by Nozaki (1972) and Pace (1986).

2.8.2 Unfolding of GroEL

The unfolding of GroEL was studied in a buffer of 50 mM-potassium phosphate pH 7.5, 20 mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol. Solutions of GroEL (final concentration 50 μg/ml) were incubated in the presence of varying concentrations of GdnHCl for 15 minutes at 20°C before any measurements of CD, fluorescence, light scattering or ATPase activity were taken. There was no further significant change on incubation for a further 45 minutes.

2.8.3 Attempted refolding of GroEL

In order to initiate the refolding process the denaturing agent was removed by dialysis at 20°C against two changes of 200 volumes of the buffer described above over a period of 6 hours. Measurements of refractive index confirmed that after dialysis the residual GdnHCl concentration was less than 0.01 M.

2.9 Characterisation of GdnHCl treated GroEL

Changes in GroEL secondary, tertiary and quaternary structure, after denaturation in the presence of GdnHCl for 15 minutes at 20°C, were determined by circular dichroism (CD), fluorescence and light scattering, respectively. The effect on ATPase activity, ANS binding and proteinase susceptibility was also studied. Parallel experiments after the removal of the denaturant by dialysis were also carried out. All experiments were carried out in the buffer described in section 2.8.2.

2.9.1 CD analysis

CD spectra were recorded at 20°C on a JASCO J-600 spectropolarimeter. Unless otherwise stated spectra were recorded in a pathlength of 1 mm. Determination

of the secondary structure content was undertaken using the CONTIN procedure at 0.2-nm intervals over the range 240-190 nm (Provencher and Glöckner, 1981). All spectra were recorded by Sharon Kelly, Department of Biological and Molecular Sciences, University of Stirling.

2.9.2 Fluorescence studies

Changes in the fluorescence of GroEL were recorded at 20°C on a Perkin Elmer LS50 spectrofluorimeter with excitation at 280 nm or 290 nm. Slit width setting was 8.0.

2.9.3 Light scattering measurements

Changes in M_r of GroEL were determined by light scattering using a Perkin Elmer LS50 spectrofluorimeter. The wavelength used was 360 nm, and the experiment carried out at 20°C. The data was analysed by the method of Parr and Hammes (1975), by using the values for refractive index increments quoted by Tashiro *et al.* (1982). Slit width setting was 2.5.

2.9.4 GroEL ATPase activity

The effect of GdnHCl on the ATPase activity of GroEL was measured by a coupled quench assay as described in section 2.4.1.

2.9.5 ANS fluorescence

Denatured and "refolded" GroEL were incubated in the presence of a 20 fold molar excess of ANS for 5 minutes at 20°C. The fluorescence of ANS in the presence and absence of protein was measured using excitation and emission wavelengths of 380 nm and 470 nm, respectively, using a Perkin Elmer LS50 spectrofluorimeter. Emission

spectra were corrected for background fluorescence due to ANS in the absence of GroEL. Slit width setting was 5.0.

2.9.6 Proteinase susceptibility

The susceptibility of GroEL to thermolysin was studied by incubating GroEL (50 μg/ml) with thermolysin (2-100% (w/w)) at 20°C for 10 minutes. The proteinase was then inactivated by the addition of EDTA to a final concentration of 2 mM, and samples analysed by SDS-PAGE.

2.10 Molecular weight determinations

SDS-PAGE was used to estimate the subunit M_r of the purified protein and proteinase treated GroEL. Unless otherwise stated, a low molecular weight (LMW) calibration kit (Sigma) was used to produce standard curves of R_f against log M_r. The standard molecular weight proteins were as given below:

Protein	Subunit M_r
Bovine serum albumin	66,000
Egg white ovalbumin	45,000
Rabbit muscle glyceraldehyde-3-phosphate dehydrogenase	36,000
Bovine erythrocyte carbonic anhydrase	29,000
Bovine pancreas trypsinogen	24,000
Soybean trypsin inhibitor	20,100
Bovine milk α–lactalbumin	14,200

2.11 Chemical modification and inactivation of GroEL ATPase activity

2.11.1 4-(iodoacetamido)salicylic acid treatment

GroEL (0.3mg/ml) was incubated in the presence of 0.5 - 2.0 mM-4-(iodoacetamido)salicylic acid in a buffer of 50 mM-triethanolamine pH 8.0 containing 10 mM-magnesium acetate, at both 4°C and 25°C. The final volume of the reaction mixture was 1 ml and at time intervals 50 µl aliquots were withdrawn and assayed for ATPase activity according to the procedure described in section 2.4.1.(a).

2.11.2 5'-p-fluorosulfonylbenzoyladenosine treatment

5'-p-fluorosulfonylbenzoyladenosine (5'pFSBA) was freshly prepared as a 10 mM stock in ethanol. GroEL (0.3 mg/ml) was incubated in the presence of 0.5 - 1.0 mM-5'pFSBA under the conditions described in section 2.11.1 and assayed for ATPase activity at various time intervals. Control experiments of GroEL in the presence of 5 - 10 % ethanol (v/v) were undertaken and assayed for ATPase activity.

2.11.3 Determination of oATP concentration

oATP was synthesised according to the procedure of Easterbrook-Smith *et al.* (1976) and the concentration determined using an extinction coefficient of 14,900 $M^{-1}cm^{-1}$ at 258nm. oATP was preincubated at 25°C for 12 hours before use.

2.11.4 oATP inactivation reaction

Inactivation was carried out by incubating GroEL (0.15 - 0.3 mg/ml) with various concentrations of oATP in 50 mM-triethanolamine pH 8.0 containing 10 mM-

magnesium acetate at 25°C. 50 µl aliquot samples were removed from the 1 ml reaction mixture at various time intervals and assayed for ATPase activity (section 2.4.1).

2.11.5 Protection by substrate against oATP inactivation

Protection experiments were carried out by including increasing concentrations of ATP in the reaction mixture described in section 2.11.4.

2.11.6 Borohydride reduction of oATP treated GroEL

The reaction of oATP with GroEL was quenched by the addition of a 10 fold molar excess of sodium borohydride over oATP. Sodium borohyride was freshly prepared as a 100 mM stock solution in distilled water containing 10 mM-NaOH and protected from light.

2.11.7 Stoichiometry of incorporation of oATP

The stoichiometry of incorporation of oATP into GroEL was studied using [^{3}H] oATP. After inactivation and sodium borohydride treatment (sections 2.11.4 and 2.11.6) the protein was precipitated, in the presence of 4 mg BSA, by the addition of 10% trichloroacetic acid (TCA) and the pellet washed exhaustively in 10% TCA to remove unbound [^{3}H] oATP. The pellet was then solubilised in 100 µl 1M-NaOH before measuring incorporated radioactivity using a Wallac 1409 Liquid Scintillation Counter (Pharmacia).

The stoichiometry of incorporation of oATP into GroEL was also studied in the presence of ATP (5 mM, final concentration), as described above.

2.12 Limited proteolysis of GroEL

GroEL (0.2 mg/ml) in a buffer of 50 mM-triethanolamine pH 8.0 containing 10 mM-magnesium acetate was digested in the presence of 2% (w/w) thermolysin at 37°C. At various time intervals aliquots were withdrawn and the reaction stopped by the addition of 2 mM-EDTA. Samples were then assayed for ATPase activity as described (section 2.4.1) and digestion monitored by analysis on 12% SDS-PAGE gels (section 2.3).

Proteolysis was also carried out in the presence of 2% (w/w) trypsin, under the conditions described above. Aliquots were withdrawn at given times and the reaction stopped by the addition of 4% (w/w) soybean trypsin inhibitor. Samples were assayed for ATPase activity and digestion monitored by SDS-PAGE (12% gels).

Limited proteolysis of GroEL by chymotrypsin (2% w/w) was carried out under the conditions described above. Reactions were stopped by boiling at 100°C for 5 minutes and the time course of digestion monitored by SDS-PAGE (12% gels).

2.13 Peptide isolation and sequencing

2.13.1 Acid-washed glassware

Glassware for protein chemistry was boiled for 1 hour in 6 N-HCl or 6 N-HNO_3 then left at room temperature overnight. It was then rinsed exhaustively in distilled water and dried in a hot oven. Glass Reacti Vials (Pierce) were used for purified peptide collection and freeze drying.

2.13.2 Preparation of oATP modified GroEL

oATP modified GroEL in the presence and absence of substrate (ATP, 5 mM final concentration) were prepared by reacting 5 mg of GroEL with 1 mM oATP in a buffer of 50 mM-triethanolamine pH 8.0 containing 10mM-magnesium acetate at 25°C. After 6 hours incubation more than 95% of the ATPase activity was lost from the sample incubated in the absence of substrate, whereas the protected sample retained 70% of its activity.

Both reaction mixtures were quenched by the addition of sodium borohydride and subjected to gel filtration on Sephadex G-25 (30 cm x 1.6 cm column), equilibrated in 0.5% ammonium bicarbonate, to remove excess reagent. The flow rate was 40 ml/hour and GroEL was detected spectrophotometrically by absorbance at 280nm. Samples were then digested with protease as described in section 2.13.3.

2.13.3 Digestion of oATP modified GroEL

oATP modified GroEL, in the presence and absence of substrate, in 0.5% ammonium bicarbonate, was digested with thermolysin, trypsin or chymotrypsin (protease : GroEL, 3% w/w) at 37°C for 24 hours. Digestion was carried out in a round bottomed acid-washed flask containing a small magnet (acid-washed) for continuous stirring. After 24 hours samples were immediately frozen in dry-ice/ethanol and then freeze dried.

2.13.4 Reverse phase HPLC of peptides

Fractionation of peptides was carried out using a C18 μBondapak reverse phase column (Waters) equilibriated in 0.15% trifluoroacetic acid in water (pH 2) and developed with a linear gradient of a mixture (70:30, v/v) of acetonitrile and 0.1%

trifluoroacetic acid. The flow rate was 1 ml/min and eluting peptides were monitored spectrophotometrically at 214nm and 260nm to detect peptide bonds and oATP modified peptides, respectively. Isolated peptide fractions were collected manually, freeze-dried and rechromatographed using a shallower linear acetonitrile gradient. Modified peptides were again collected manually, freeze dried and stored at -20°C.

2.13.5 Modified peptide sequence analysis

The modified GroEL peptide was sequenced on an Applied Biosystems 477A pulsed liquid phase sequencer with on line detection of amino acid thiohydantoins by a model 120A analyser. The instrument was operated by Dr. B. Dunbar at the SERC Protein Sequencing Facility, University of Aberdeen.

2.13.6 Mass spectroscopy of the modified peptide

The GroEL modified peptide was analysed on a VG BioTech Platform single quadropole mass spectrometer in the electrospray ionisation mode. An aliquot of the sample was diluted 1:1 with acetonitrile and was injected into a mobile phase of 1:1 acetonitrile:water flowing at 5µl / min via a 10µl loop. The mass spectrometer was scanned over the range 300-1800 in 10 seconds and the data was acquired in multi-channel analysis mode. Spectra were acquired in both positive and negative ion mode and the mass spectrometer was calibrated from a separate introduction of a mixture of polyethylene glycols. The instrument was operated by Dr. H. Major, VG BioTech, Macclesfield,Cheshire, U.K.

2.14 Characterisation of proteinase K-treated GroEL

2.14.1 Limited proteolysis with proteinase K

GroEL (0.1 mg/ml) in 50 mM-triethanolamine pH 8.0 containing 10 mM-magnesium acetate was digested in the presence of proteinase K (2% w/w) at 25°C. At time intervals aliquots were removed and the reaction stopped by the addition of 2 mM-PMSF for 5 minutes. Digestion was follwed on 12% SDS-PAGE gels and samples were assayed for ATPase activity.

GroEL incubated in the presence of 1-mM oATP or 5mM-ATP was also subjected to proteinase K digestion, as described above.

2.14.2 Electroblotting of proteinase K-treated GroEL

GroEL was digested and run on SDS-PAGE as described in section 2.14.1. Proteins were then transferred onto polyvinyl difluoride (PVDF) membranes using a modification of the method of Matsudaira (1987), described below.

The blotting was carried out in 10mM CAPS buffer pH 11.0 (titrated with 2 M NaOH), which contains no free amino groups to interfere with the sequencing.

The PVDF membrane was pre-treated by soaking in 100% AnalaR methanol for a few seconds, then 50% methanol/water before transferring to electroblotting buffer (10 mM CAPS pH 11.0).

The gel was removed from the electrophoresis apparatus and soaked in 2 x 250 ml of electroblotting buffer, along with 2 pieces of Whatman 3MM filter paper cut to gel size with a 2 x 2 cm overlap.

The PVDF membrane was placed over the gel then sandwiched between the two pieces of filter paper and ensuring that no air was trapped, placed in the electroblotting apparatus. Electroblotting was carried out at 50V at room temperature for 90 minutes.

The PVDF membrane was then rinsed with deionised water for 5 minutes, then dipped into 100% methanol for a few seconds prior to staining (section 2.14.3).

2.14.3 Staining and destaining of PVDF membranes

PVDF membranes were stained for 2 minutes in 0.1% (w/v) Coomassie Blue R-250 in 50% (v/v) methanol and destained with several changes of 50% (v/v) methanol / 10% (v/v) acetic acid. Membranes were then rinsed with several changes of deionised water and allowed to air dry before storing at - 20°C.

2.14.4 N-terminal sequence analysis

Protein blotted onto PVDF membranes were sequenced by Dr B. Dunbar, University of Aberdeen, as described in section 2.13.5.

2.14.5 Stoichiometry of oATP incorporation into proteinase K-treated GroEL

GroEL (0.2 mg/ml) was digested by proteinase K (2% w/w) as described in section 2.14.1, for 1 hour before stopping the reaction by the addition of 2 mM-PMSF. The samples were then incubated in the presence of 0.6 mM-[^{3}H] oATP, in the presence or absence of 5 mM-ATP, at 25°C for 6 hours before treatment with sodium borohydride and precipitation of the protein as described in section 2.11.7. Incorporation of radioactivity was measured using a Wallac 1409 Liquid Scintillation Counter (Pharmacia).

2.15 Titration of thiol groups with DTNB

GroEL (0.2mg/ml) was reacted with 250 μM-5,5'-dithiobis(2-nitrobenzoate) (DTNB) in 50 mM-sodium phosphate buffer pH8.0 at 25°C in a final volume of 1ml. The reaction was initiated by addition of 0.1% SDS. The absorption at 412nm was measured against a blank lacking GroEL on a Gilford/Unicam model 252 uv/vis spectrophotometer. The formation of TNB was calculated using a molar extinction coefficient of 13 600 $M^{-1}cm^{-1}$ (Ellman, 1959).

Free titratable GroEL thiol groups in the presence of 1 mM-oATP, with or without 5 mM-ATP, in the buffer described above were also measured, according to the method described.

Chapter 3 Purification of GroEL

3.1 Introduction

This chapter describes the purification of GroEL from the *E.coli* strain DH1 carrying either the plasmid pGT3270 or pND5. The purification was monitored by SDS-PAGE since the ATPase activity of GroEL is too weak and non specific to monitor. The purification was modified from that described previously (Hendrix, 1979; Chandrasekhar *et al.*, 1986) in order to achieve the high degree of purity required for both the chemical modification (chapter 4) and unfolding/refolding experiments (chapter 5).

Purification of GroEL to a high degree of homogeneity can be complicated by the presence of bound 'substrate' polypeptides associated with GroEL, which can vary with preparations (Hayer-Hartl and Hartl, 1993). The method described in this chapter utilizes a hydrophobic interaction chromatography step, not previously reported, in order to yield electrophoretically homogenous GroEL.

3.2 GroEL purification.

Unless otherwise stated all manipulations following cell breakage were performed at 4°C.

Step 1: Extraction and centrifugation

A 20g batch (wet weight) of *E.coli* cells were broken as described in section 2.5.4. Following cell breakage DNAse was added to a final concentration of 1 mM and the lysate stirred slowly on ice for 30 minutes. The extract was then centrifuged for 45 minutes at 25,000 g using a Beckman J2-HS centrifuge with a JA-20 fixed angle rotor. GroEL was purified from the resulting cell-free extract.

Step 2: 'Negative' chromatography on DEAE-cellulose

The crude extract was pumped at 60 ml/h through a column (7.5 cm x 5 cm) of DE52 pre-equilibrated with 50 mM-trisHCl pH 7.5 containing 0.1 M-$(NH_4)_2SO_4$, 1 mM-EDTA, 10% glycerol, 5 mM-2-mercaptoethanol, 0.05% triton X-100. The column was washed with equilibration buffer and the run through (brown colour), typically 150 ml, collected.

Step 3: Salt fractionation with $(NH_4)_2SO_4$

The run through from the DE52 column was brought to 50% saturation by addition of solid $(NH_4)_2S0_4$ (314 g / l). After stirring for 30 minutes, the precipitated proteins were collected by centrifugation at 25,000 g for 30 minutes using a Beckman J2-HS centrifuge with a JA-20 fixed angle rotor.

Step 4: Sephacryl S-200 gel permeation chromatography

Precipitated protein was resuspended in 4 ml S-200 buffer (50 mM-trisHCl pH 8.0 containing 5 mM-EDTA, 10% glycerol, 5 mM-2-mercaptoethanol, 1 M-KCl, 0.5% triton X-100) and loaded onto a Sephacryl S-200 column (120 cm x 3 cm) equilibrated in the same buffer. The flow rate was 10 ml/h and GroEL containing fractions, eluting in the void volume of the column, were identified by 12% PhastGel SDS-PAGE as shown in Fig. 3.1. GroEL-containing fractions were pooled together and dialysed overnight versus 100 volumes of 20 mM-trisHCl pH 7.5 containing 5 mM-EDTA.

Step 5: Chromatography on Mono Q

The GroEL containing-pool was applied to a Mono Q HR 5/5 anion exchange column equilibrated in 20 mM-trisHCl pH 7.5 containing 5 mM-EDTA. The column

was eluted with a 25 ml 0-1 M linear NaCl gradient in the buffer described above; the GroEL protein eluted at 0.45 M-NaCl as shown in Fig. 3.2. The flow rate was 1 ml/min and 0.5 ml fractions were collected. Fractions containing GroEL (identified by PhastGel SDS-PAGE) were pooled and dialysed against 100 volumes of 50 mM-potassium phosphate buffer pH 7.0 containing 1.5 M-$(NH_4)_2SO_4$.

Step 6: Chromatography on Alkyl-Superose

The partially pure GroEL (85-90% pure) was then loaded onto an Alkyl-Superose HR 5/5 hydrophobic interaction chromatography column, equilibrated in 50 mM-potassium phosphate pH 7.0, 1.5 M-$(NH_4)_2SO_4$. GroEL was eluted with a linear 1.5-0 M-$(NH_4)_2SO_4$ gradient in a volume of 20 mls. GroEL was eluted at approximately 0.6 M-$(NH_4)_2SO_4$ (Fig. 3.3). The flow rate was 0.5 ml/min and 0.5 ml fractions were collected. GroEL was normally judged to be greater than 95% pure on SDS-PAGE, although occasionally a further purification step was required (step 7).

Step 7: Superose 6 chromatography

As a final purification step the pooled GroEL-containing fractions were dialysed against 50 mM-trisHCl pH 7.5 containing 1mM-EDTA and 0.2 M-NaCl and concentrated by ultrafiltration using Amicon Centriprep 30 cells. 0.5 ml samples were applied to a Superose 6 HR 10/30 column and eluted with the same buffer. The flow rate was 0.5 ml/min and 0.5 ml fractions were collected. The elution profile of GroEL is shown in Fig. 3.4.

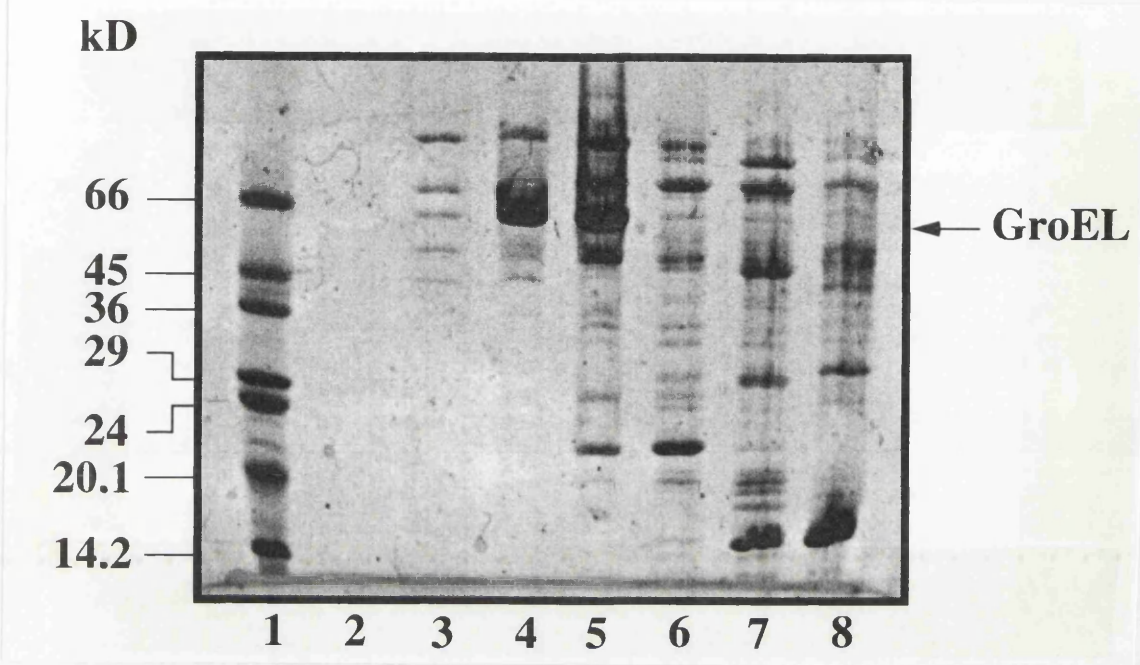


Figure 3.1 SDS-PAGE analysis of GroEL eluting at Sephacryl S-200 void volume

Fractions eluting near the void volume of the S-200 column were analysed under denaturing conditions on a 12.5% Phast Gel as described in section 2.3 and stained for protein using Coomassie blue as described in section 2.3.1.

Lane 1: markers; BSA, ovalbumin, GAPDH, carbonic anhydrase, trypsinogen, trypsin inhibitor, α-lactalbumin.

Lanes 2-8: material eluting near the S-200 void volume. The position of GroEL is indicated by the arrowhead.

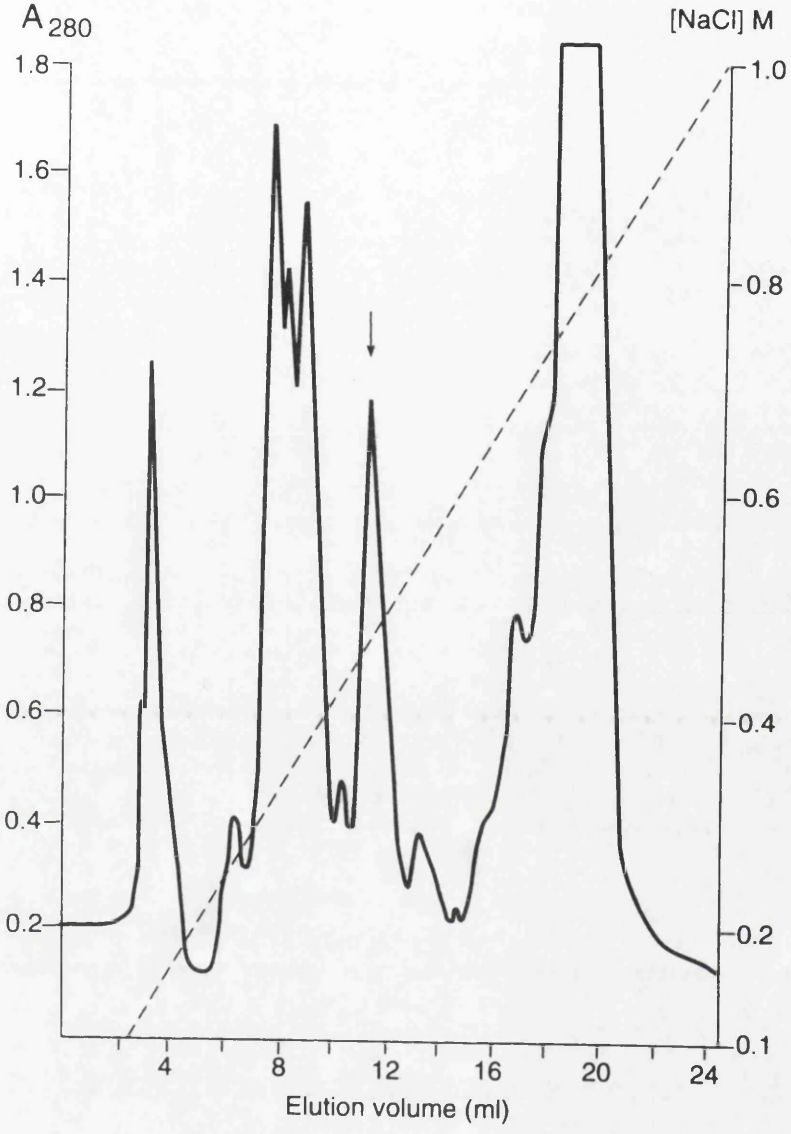


Figure 3.2 Chromatography on Mono Q

GroEL-containing pool from Sephacryl S-200 chromatography was loaded onto a Mono Q column and eluted with an increasing salt gradient (indicated by the *dashed line*) as described in section 3.2. The protein elution profile was measured by absorbance at 280 nm (A_{280}), with the position of GroEL elution indicated by the arrowhead.

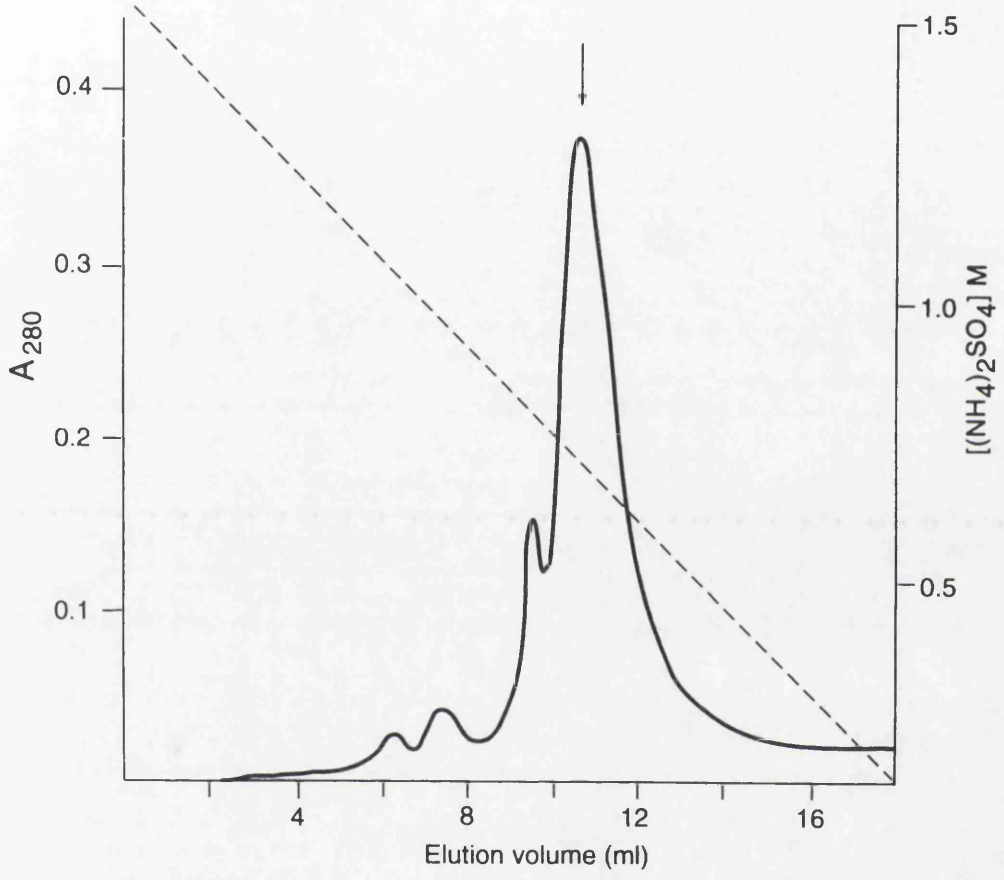


Figure 3.3 Chromatography on Alkyl-Superose

GroEL from chromatography on Mono Q was loaded onto an Alkyl-Superose column and eluted with a decreasing salt gradient (indicated by the *dashed line*) as described in section 3.2. The protein elution profile was measured by absorbance at 280 nm (A_{280}), with the position of GroEL elution indicated by the arrowhead.

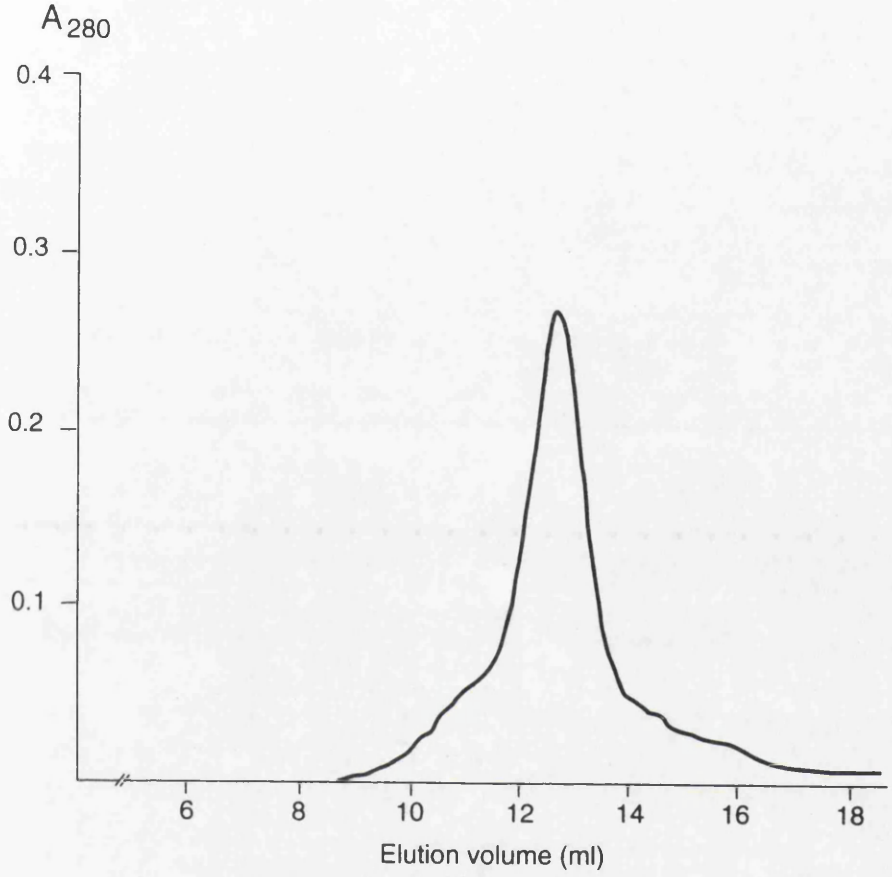


Figure 3.4 Chromatography on Superose 6

GroEL from chromatography on Alkyl-Superose was loaded onto a Superose 6 column as described in Section 3.2. The protein elution profile was measured by absorbance at 280nm (A_{280}).

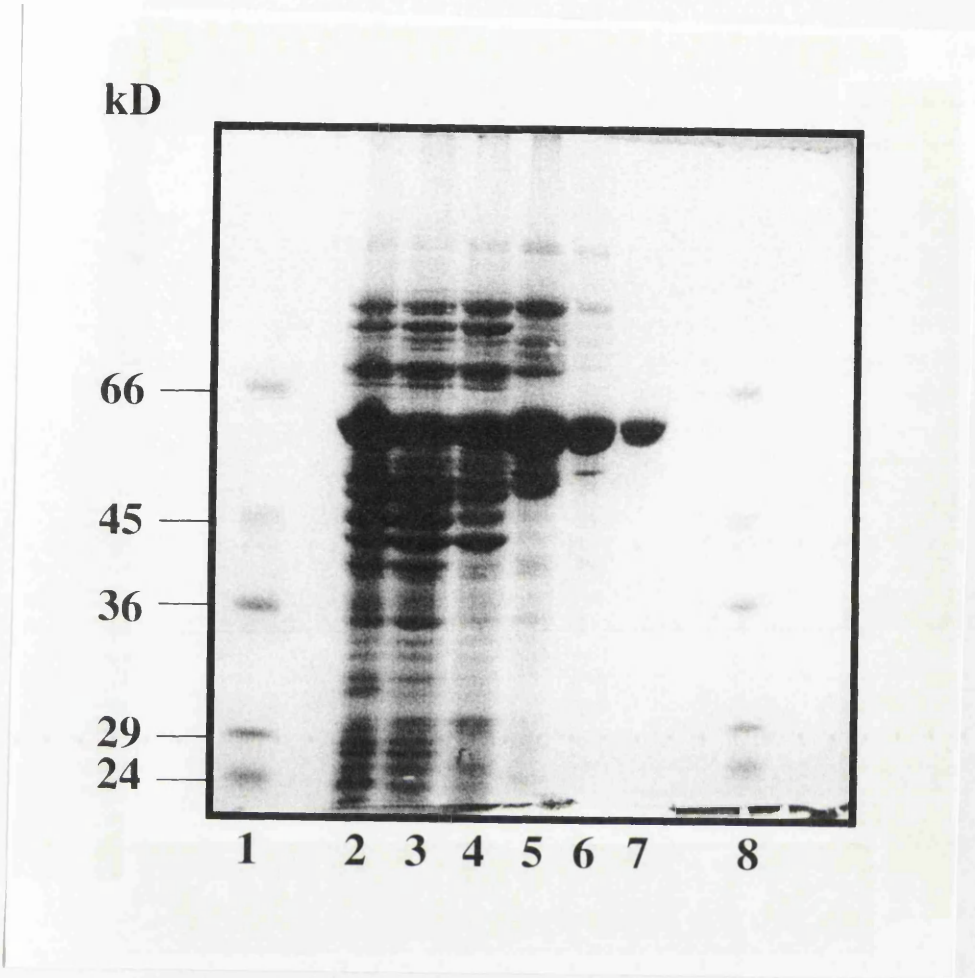


Figure 3.5 SDS-PAGE analysis of purified GroEL

Samples from each step of the purification were analysed on a 12% denaturing gel as described in section 2.3 and stained for protein using Coomassie blue as described in section 2.3.1.

Lanes 1 and 8: markers; BSA, ovalbumin, GAPDH, carbonic anhydrase, trypsinogen (section 2.10)

Lanes 2-7: crude extract, $(NH_4)_2SO_4$ cut, DE52 pool, Sephacryl S-200 pool, Mono Q pool and Alkyl-Superose pool respectively.

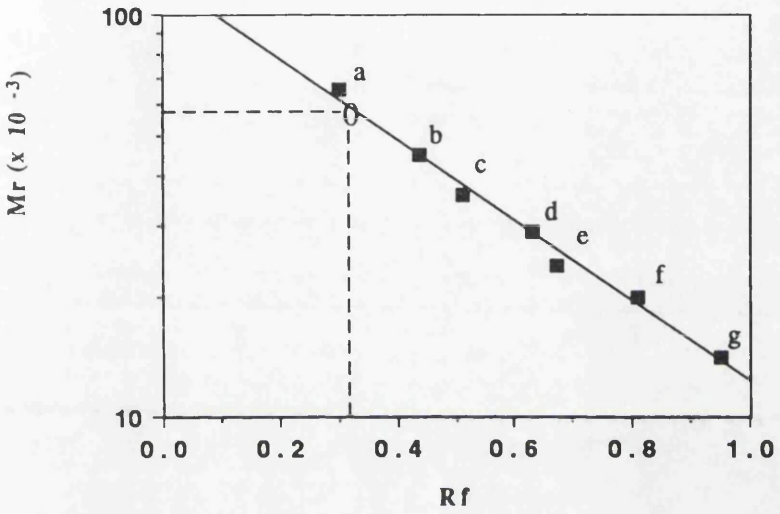


Fig. 3.6 Molecular mass of GroEL

Purified GroEL was subjected to SDS-PAGE (12%) as described in section 2.3. The R_f values of standard proteins (section 2.10) were calculated and plotted against the log M_r. The R_f of the purified GroEL is shown (0). The standard marker proteins are ; a = BSA (66,000 M_r), b = ovalbumin (45,000 M_r), c = GAPDH (36,000 M_r), d = carbonic anhydrase (29,000 M_r), e = trypsinogen (24,000 M_r), f = trypsin inhibitor (20,100 M_r) and g = α-lactalbumin (14,200 M_r).

3.3 Purity of GroEL

Only one band was visible on SDS-PAGE (12%) after staining with Coomassie blue (Fig. 3.5, lane 7) after purification, as described above. GroEL was judged to be over 95% homogenous by laser densitometric analysis on a 12% SDS-PAGE gel.

3.4 Molecular weight determination of GroEL

The subunit M_r of GroEL was 57,400, as determined by SDS-PAGE (12%) using Sigma low molecular weight markers (Fig. 3.5). This value is in close agreement with the subunit M_r of 57, 259 as derived from the amino acid sequence of GroEL (Hemmingsen *et al.*, 1988).

3.5 Purification yield of GroEL

A typical yield of pure GroEL from 20g (wet weight) of *E.coli* DH1/pGT3270 was in the order of 15 - 20 mg per preparation. The strain carrying plasmid pND5, which contains surplus genes over plasmid pGT3270, was found to be less efficient for purification, with the plasmid seemingly regressing.

Purified GroEL was stable for at least 2 months when stored at - 20°C in 50% glycerol (section 2.6.1).

3.6 Discussion

A highly homogenous preparation of GroEL could be obtained using the purification protocol described above. Most published protocols are modifications of the methods of Hendrix (1979) and Chandrasekhar *et al.* (1986) and as such the final purification step is generally the anion exchange chromatography step (e.g. Buchner *et*

al.,1991; Fisher, 1992). Purification of GroEL from *E.coli* DH1 (with either plasmid pND5 or pGT3270) required additional chromatography step(s) involving hydrophobic interaction chromatography and / or gel filtration.

After the Mono Q step GroEL was still not pure as judged by SDS-PAGE (Fig. 3.5, lane 6) with several contaminating proteins still present. This is not all that surprising when considering that the role of GroEL *in vivo* is to interact with many different polypeptides to assist in their assembly into native structure. Thus at least some of these contaminants may be actually associated with the GroEL protein itself and thus difficult to dissociate (Hayer-Hartl and Hartl, 1993). It should also be noted that different *E.coli* strains as well as plasmids were used in this work compared to the published methods mentioned above.

The use of the Alkyl-Superose column was generally sufficient for the final purification step (Fig. 3.5, lane 7). GroEL was able to bind to Alkyl-Superose presumably via exposed hydrophobic patches on the surface of the protein. This is consistent with reports that the hydrophobic probes ANS (Mendoza *et al.*,1992; chapter 5) and bisANS (Mendoza *et al.*,1991) showed enhanced fluorescence when interacting with GroEL. Release of GroEL from the column was achieved by a decreasing salt gradient to yield electrophoretically homogenous protein. Thus the use of hydrophobic interaction chromatography can be useful for final purification of GroEL, possibly by removing associated polypeptides from GroEL which themselves are binding through hydrophobic interactions. The purified GroEL was used for the experiments outlined in chapters 4 and 5.

Chapter 4 Affinity labelling the GroEL ATPase active site

4.1 Introduction

The 3-dimensional arrangement of the polypeptide chain in GroEL has, as yet, still to be elucidated and as such the amino acid residues of GroEL involved in ATP binding and hydrolysis have not yet been identified; this chapter deals with the investigation of such residues.

Many ATP binding proteins posses sequence motifs such as the type A motif (Walker *et al.*, 1982), often abbreviated to -G-$(X)_4$-G-K-T/S-. Although some members of the chaperonin 60 class show the presence of a putative 'extended' type A motif containing an extra two amino acids between the first and second glycine residues of the consensus sequence (Tsugeki *et al.*, 1992), the reported cDNA derived amino acid sequence of GroEL (Hemmingsen *et al.*, 1988) does not contain any such sequence. In order to define amino acid residues involved in ATP binding an affinity labelling approach was undertaken.

4.2 Affinity labelling with oATP

4.2.1 Background

Although chemical modification using group specific reagents has proved to be a useful tool in identifying amino acid residues involved at the active site of proteins, a potentially more powerful technique in defining such residues is by affinity labelling of the active site. This method can avoid some of the non-specific labelling often observed with group specific reagents. Specifically designing reagents, with structural similarity to the natural substrate, which contain reactive functional groups has been widely used in investigating purine nucleotide sites in proteins (Colman, 1983; Easterbrook-Smith *et al.*, 1976; Roy and Colman, 1980).

One such reagent is 2', 3' dialdehyde ATP (oATP) which has been successfully used to specifically label a number of ATP binding proteins (Easterbrook-Smith *et al.*, 1976; King and Carlsson, 1981; Kumar *et al.*,1979). oATP is structurally similar to the natural nucleotide, containing the purine ring and the correct number of phosphate groups; however the ribose ring has been perturbed by periodate oxidation of the 2',3'-*cis* -diol to yield the corresponding dialdehyde (Fig. 4.1). The general mechanism of reaction is proposed to be the formation of a Schiff base by the nucleophilic attack of one of the dialdehyde groups by the α-amino group of lysine, although additional non-Schiff base forming reactions have been reported. The results of treating GroEL with oATP are reported in this chapter. Other potential active site reagents used are also described.

4.2.2 Criteria for a potential affinity label

The following criteria should be fulfilled when designating a compound as an affinity label;

(a) Kinetic evidence for the initial formation of a reversible enzyme-inhibitor complex prior to irreversible inactivation should be demonstrated.

(b) The extent of modification by the affinity label should be more limited than that produced by a structurally unrelated reagent with the same reactive group.

(c) The presence of the natural substrate that binds to the same site should decrease the rate of inactivation by the affinity label.

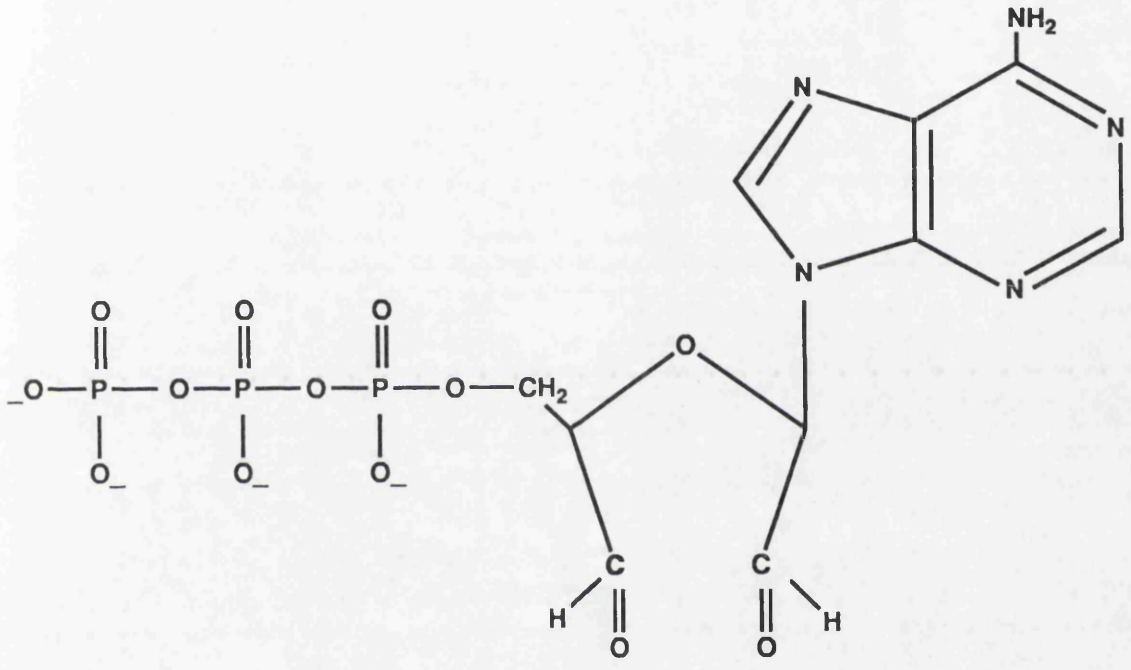


Fig. 4.1 2', 3' dialdehyde ATP (oATP)

4.2.3 Kinetics of oATP inactivation of GroEL ATPase activity

Incubation of GroEL with oATP at 25°C in 50 mM-TEA, 10 mM-magnesium acetate, pH 8.0 resulted in a loss of ATPase activity. In control experiments, in the absence of oATP, no loss in activity was observed. The fraction of activity remaining was calculated as the percentage of activity remaining (A/A_0, where A is the activity at each time and A_0 is the initial rate of activity, *i.e.* at time 0) at a given time (t). The plot of the logarithm of A/A_0 against t at various concentrations of oATP shows pseudo first-order kinetics (Fig. 4.2 A). The rate of inactivation is dependent on oATP concentration. Table 4.1 shows the half life period ($t_{1/2}$)of inactivation determined at various oATP concentrations, calculated from the pseudo first order plot. Inactivation was unaffected whether or not the reaction was quenched by sodium borohydride.

Table 4.1

Half-life ($t_{1/2}$) and first order rate constant (k_{obs}) for the inactivation of GroEL ATPase activity at different oATP concentrations

[oATP] mM	$t_{1/2}$ (min)	k_{obs} / min
0.10	158	0.0044
0.25	63.5	0.0109
0.50	38.0	0.0182
1.00	22.2	0.0312
2.50	14.0	0.0495

The half-life period ($t_{1/2}$) for the inactivation process is the time after which the activity has decreased to half of its original activity, *i.e.* $A = A_0/2$. The half life period may be expressed as

$$t_{1/2} = \text{In } 2/k_{obs} = 2.303 \log 2/k_{obs} = 0.693/k_{obs}$$

From this relationship the pseudo first-order rate constant for inactivation, $k_{obs} = 0.693/t_{1/2}$, was calculated.

The relationship between k_{obs} and desired constants (*i.e.* K_I) may be described by the equation (Kitz and Wilson, 1962)

$$k_{obs} = k_2 / [(K_I / [I]) + 1] \qquad \text{[equation (1)]}$$

This describes the observed rate constant and is derived from Michaelis-Menten type derivations of the following system

$$E + I \underset{k\text{-}1}{\overset{k1}{\rightleftharpoons}} E\text{-}I \xrightarrow{k2} EX$$

where E represents the free enzyme, I, the inhibiting reagent, E - I the reversible enzyme inhibitor complex and EX, the inactivated enzyme. K_I is the concentration of I giving the half-maximum inactivation rate, i.e. $K_I = (k_{-1} + k_2)/k_1$.

A double reciprocal plot of k_{obs} versus oATP concentration (Kitz and Wilson, 1962) (Fig 4.2B) showed saturation kinetics with the pseudo first-order rate constants proportional to the concentration of oATP. This is indicative of the formation of a reversible GroEL-oATP complex prior to irreversible inactivation, one of the criteria for an affinity label. In accordance with equation (1) the apparent K_I determined from

Fig 4.2 Kinetics of inactivation of GroEL ATPase with oATP

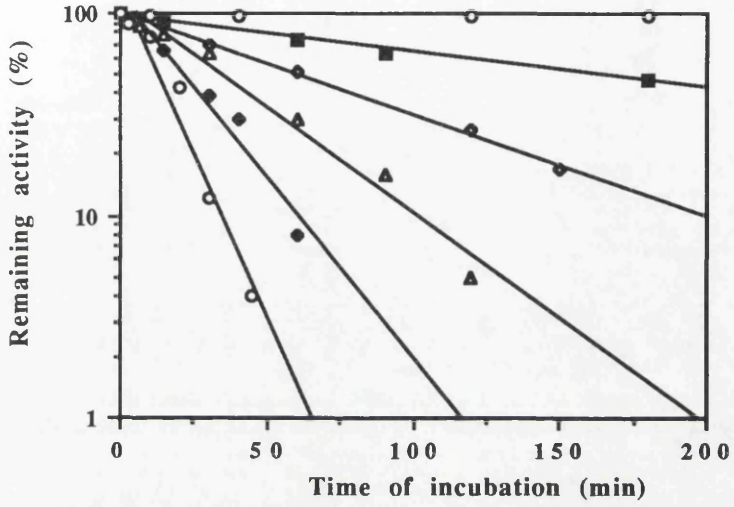
A. Pseudo first-order plots for inactivation

GroEL (~ 3.5 μM per subunit) was incubated with increasing concentrations of oATP in 50 mM-TEA, 10 mM-magnesium acetate, pH 8.0 at 25°C. At time intervals aliquots were removed and assayed for ATPase activity as described in section 2.4.1 (a). The concentrations of oATP used were , 0.0 mM (o), 0.1 mM (■), 0.25 mM (◊), 0.5 mM (Δ), 1.0 mM (•) and 2.5 mM (o).

B. Determination of K_I and k_2

Pseudo first-order rate constants (k_{obs}) calculated in part A were proportional to oATP concentration in this double reciprocal plot. The second order rate constant for the inactivation of GroEL ATPase activity by oATP, as denoted by the slope of the curve (K_I / k_2), is 46 M^{-1} min^{-1}.

A


100
10
1
Remaining activity (%)
0
50
100
150
200
Time of incubation (min)

B

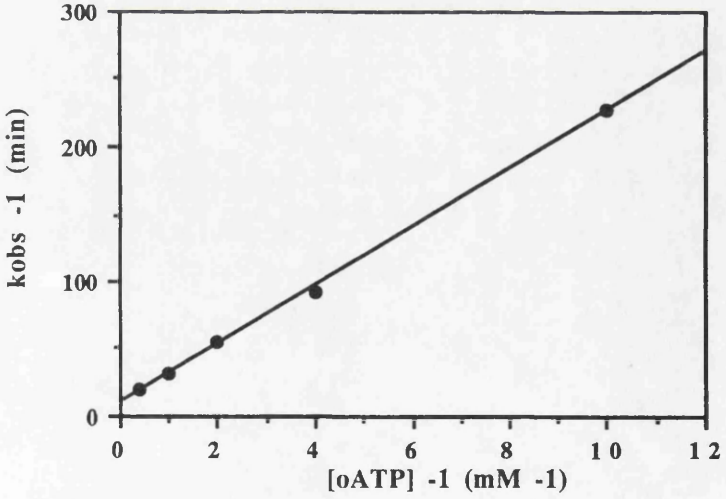

300
200
100
0
kobs -1 (min)
0
2
4
6
8
10
12
[oATP] -1 (mM -1)

Fig. 4.2B was 2.11 mM, with the maximum rate of inactivation (k_2) at saturating concentrations of oATP was 0.1 min^{-1}.

4.2.4 Substrate protection against oATP inactivation

Substrate protection of an enzyme is one of the criteria for an affinity label. Protection of substrate against inactivation strengthens the argument that inactivation is likely to be due to a reaction at or near the active site (Colman, 1983).

The ATPase activity of GroEL could be protected against inactivation by the prior inclusion of ATP (Fig. 4.3A). The data from the pseudo first-order plot indicates that the half-life of inactivation increases with ATP concentration (table 4.2). Increased protection against 1 mM-oATP inactivation was observed with increasing ATP concentrations, with up to 70% of the original activity remaining in the presence of 5 mM-ATP (Fig. 4.3B). Therefore it appears as though oATP is interacting with critical amino acids at or near the ATP binding site.

Table 4.2

Effect of substrate concentration on oATP inactivation of GroEL ATPase activity

[ATP] mM	$t_{1/2}$ (min)	k_{obs}/min
0.0	22.0	0.0315
1.0	48.5	0.0143
5.0	91.0	0.0076

Fig.4.3 Substrate protection against oATP inactivation of GroEL ATPase activity

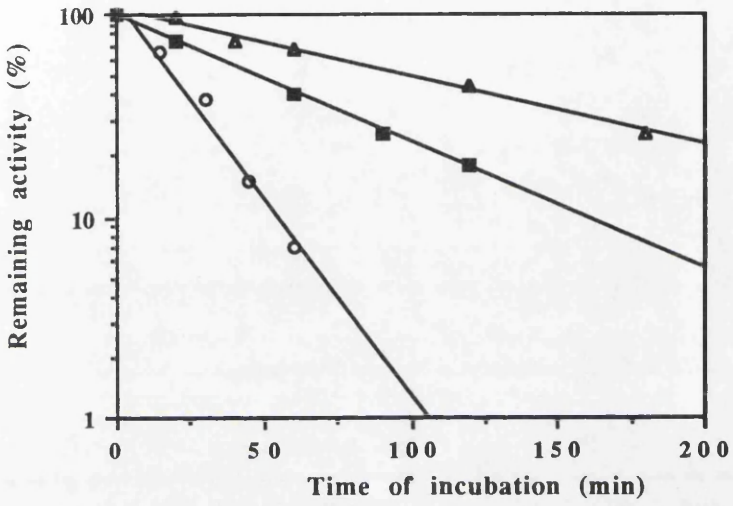
A. Pseudo first-order plots for inactivation

GroEL (~ 3.5 μM per subunit) was incubated in the presence of 1 mM-oATP and increasing concentrations of ATP in 50 mM-TEA, 10 mM-magnesium acetate, pH 8.0 at 25°C. At time intervals aliquots were removed and assayed for ATPase activity as described in section 2.4.1(a). The concentrations of ATP present were 0 mM (o), 1 mM (■) and 5 mM (Δ).

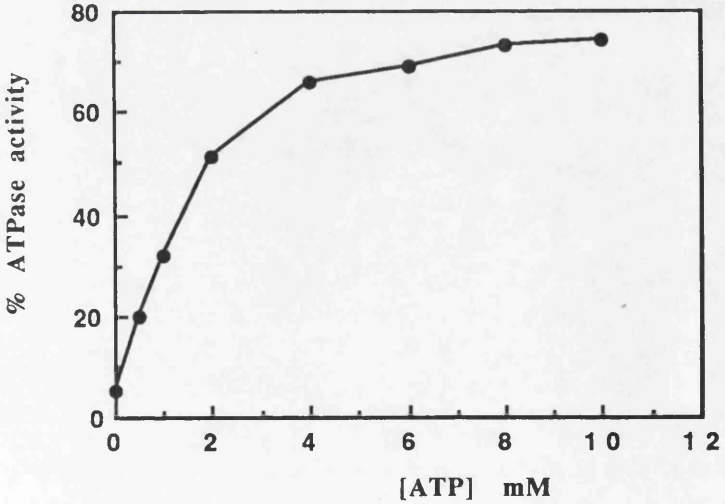
B. Protection against oATP inactivation by inclusion of ATP

GroEL (~ 5.3 μM per subunit) was incubated in the presence of 1 mM oATP and varying concentrations of ATP in a buffer of 50 mM-TEA, 10 mM-magnesium acetate, pH 8.0 at 25°C. After 6 hours incubation samples were withdrawn and assayed for ATPase activity as described in section 2.4.1(a).

A


Remaining activity (%)
100
10
1
0
50
100
150
200
Time of incubation (min)

B


% ATPase activity
80
60
40
20
0
0
2
4
6
8
10
12
[ATP] mM

4.2.5 Stoichiometry of incorporation of oATP into GroEL

In order to measure the incorporation of oATP into GroEL, the protein was inactivated by [2', 5', 8 - ^{3}H] oATP. Results of a number of independent experiments are presented in Figure 4.4, where the inactivation process is directly related to the increasing incorporation of [2', 5', 8 - ^{3}H] oATP. When the enzyme is completely inactivated this corresponds to the incorporation of approximately 1.30 moles of oATP per active site. Therefore although additional reactions appear to occur at other sites complete inactivation of the GroEL ATPase activity is due to the incorporation of about 1 oATP per subunit.

4.2.6 Stoichiometry of incorporation of oATP into substrate protected GroEL

Incorporation of oATP into GroEL in the presence of ATP (section 2.11.7) was reduced compared to incorporation in the absence of substrate (table 4.3). Substrate-protected GroEL retained 65% of its ATPase activity corresponding to the incorporation of around 0.6 moles of oATP per GroEL monomer. This is directly comparable to the incorporation of approximately 1.6 moles of oATP in inactivated GroEL in the absence of ATP (table 4.3). This suggests that the amino acid residue(s) recognised by oATP are no longer available for modification in the presence of substrate. Thus by the criteria of protection by substrate and stoichiometry of modification it appears that oATP acts as an affinity label of the GroEL ATPase site.

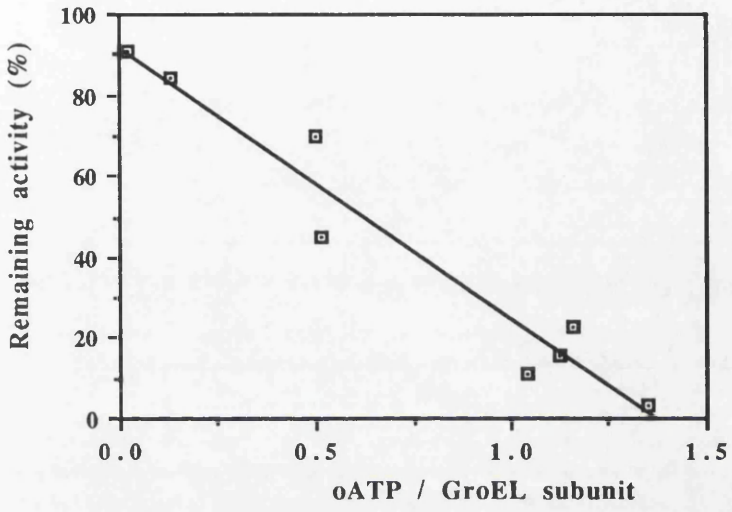


Fig. 4.4 Stoichiometry of inactivation of GroEL ATPase activity by oATP

GroEL (~ 3.5 μM per subunit) was incubated in the presence of 0.6 mM-[2', 5', 8 - 3H] oATP in 50 mM-TEA, 10 mM-magnesium acetate, pH 8.0. At time intervals aliquots were removed and the reaction quenched by a 10 fold excess of sodium borohydride for 30 minutes (section 2.11.6). Samples were then assayed for ATPase activity (section 2.4.1(a)) before measuring incorporated radioactivity (section 2.11.7).

Table 4.3

Stoichiometry of inactivation of GroEL ATPase activity by oATP

Incubation	Activity (%)	oATP/monomer
GroEL*	100	0.0
GroEL*/ 1 mM-oATP	5	1.57 ± 0.64
GroEL*/ 1 mM-oATP / 5 mM-ATP	65	0.62 ± 0.22

* GroEL concentration was 5 μM per subunit

4.3 Identification of the oATP modified peptide

The evidence provided in the previous sections (section 4.2) suggested that oATP is indeed an affinity label acting at the GroEL ATPase active site, and that inactivation is due to the incorporation of around 1 oATP per GroEL subunit. In order to identify the site of modification a differential peptide mapping approach was employed. This technique allows identification of regions of the protein which are modified by a particular site directed reagent but are protected in the presence of substrate. Adenine containing ligands absorb strongly at 260 nm and therefore it was not necessary to use radiolabelled oATP in order to detect the oATP modified peptide.

4.4 Peptide mapping

4.4.1 Limited proteolysis of GroEL

In order to determine the susceptibility of GroEL to proteolysis, GroEL was digested in the presence of 2% w/w chymotrypsin, trypsin or thermolysin under the

conditions described in section 2.12. Complete proteolysis was observed after 24 hours digestion with all three proteases as analysed by SDS-PAGE and confirmed by reverse phase chromatography using a Waters μBondapak C18 column.

The time course of proteolysis of GroEL in the presence of 2% w/w thermolysin, as analysed by SDS-PAGE, is shown in Figure 4.5. Thermolysin was selected for peptide mapping of oATP modified GroEL.

4.4.2 Proteolysis of oATP modified GroEL

GroEL (~50 μM per subunit) was inactivated in the presence of 1 mM-oATP for 6 hours by which time less than 5% of the ATPase activity remained. A parallel experiment was performed in the presence of 5 mM-ATP (section 2.13.2). Both reaction mixtures were quenched by a 10 fold excess of sodium borohydride and subjected to gel filtration on Sephadex G-25 to remove excess oATP. The protein was then transferred to acid-washed round bottomed flasks and digested for 24 hours in the presence of 3% w/w thermolysin at 37°C in 0.5% ammonium bicarbonate, with continuous stirring. After 24 hours samples were removed and immediately frozen with dry-ice in ethanol before lyophilization. Samples were stored at -20°C until required for peptide fractionation.

4.4.3 Differential peptide mapping of oATP modified-GroEL

oATP modified-GroEL in the presence and absence of ATP was prepared and proteolysed as described above. Freeze-dried samples were resuspended in deionised distilled water and fractionated by reverse phase HPLC using a Waters μBondapak C18 column. Peptides were eluted using a linear acetonitrile gradient (section 2.13.4) and monitored at both 214 nm (to detect peptide bonds) and at 260 nm (in order to detect adenine-containing peptides). When monitored at 214 nm the complex peptide profiles

were sufficiently similar to make identification of any modified peptide very difficult (Fig. 4.6). The presence of an additional peptide species seemed to elute after approximately 32 minutes in the modified sample (Fig. 4.6A). The traces monitored at 260 nm , however, showed one region which was substantially protected by ATP (Fig. 4.7). This peptide, clearly eluting after approximately 32 minutes (25% CH_3CN), was present in the modified and inactivated GroEL (Fig. 4.7A) but absent in the protected sample (Fig. 4.7B). Other differences in regions of the peptide profiles were much less pronounced. The difference in profiles observed at approximately 32 minutes were highly reproducible between separate experiments.

4.4.4 Isolation and purification of the modified peptide

The material eluting at approximately 32 minutes was collected and re-chromatographed by reverse phase HPLC under the same conditions described previously (section 2.13.4). The elution profile, monitored at 214 nm, showed a major component emerging at the same point in the gradient, but which contained a pronounced shoulder (Fig. 4.8A).

In an attempt to purify this material further aliquots of the sample were loaded onto a Pierce microbore (1 x 100 mm) C8 column and eluted with a shallow acetonitrile gradient, similar to what was used previously. Again, the peak eluting was not symmetrical and, as before - whether monitored at 260 nm or 214 nm (Fig. 4.8B) - showed evidence of a shoulder, indicating that more than one peptide species was present. This material was, however, collected for amino acid sequence analysis.

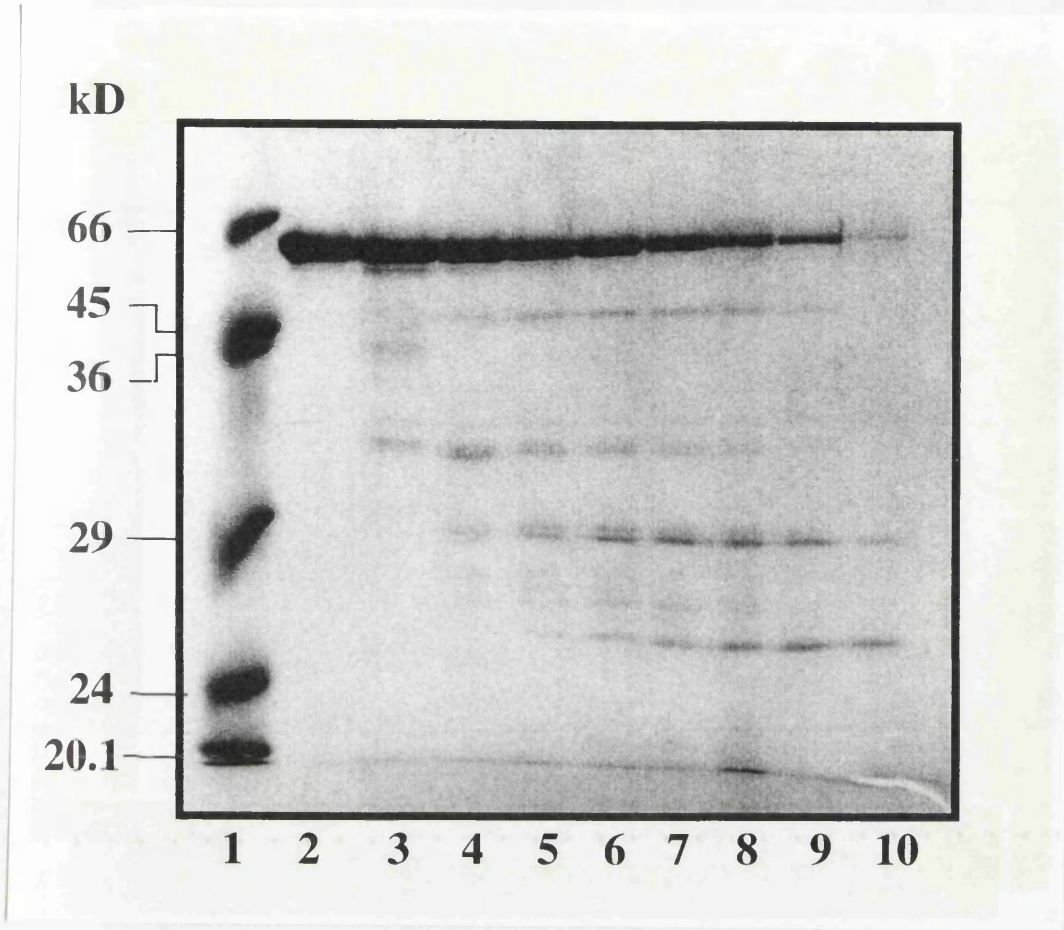


Fig. 4.5 SDS-PAGE analysis of thermolysin limited proteolysis of GroEL

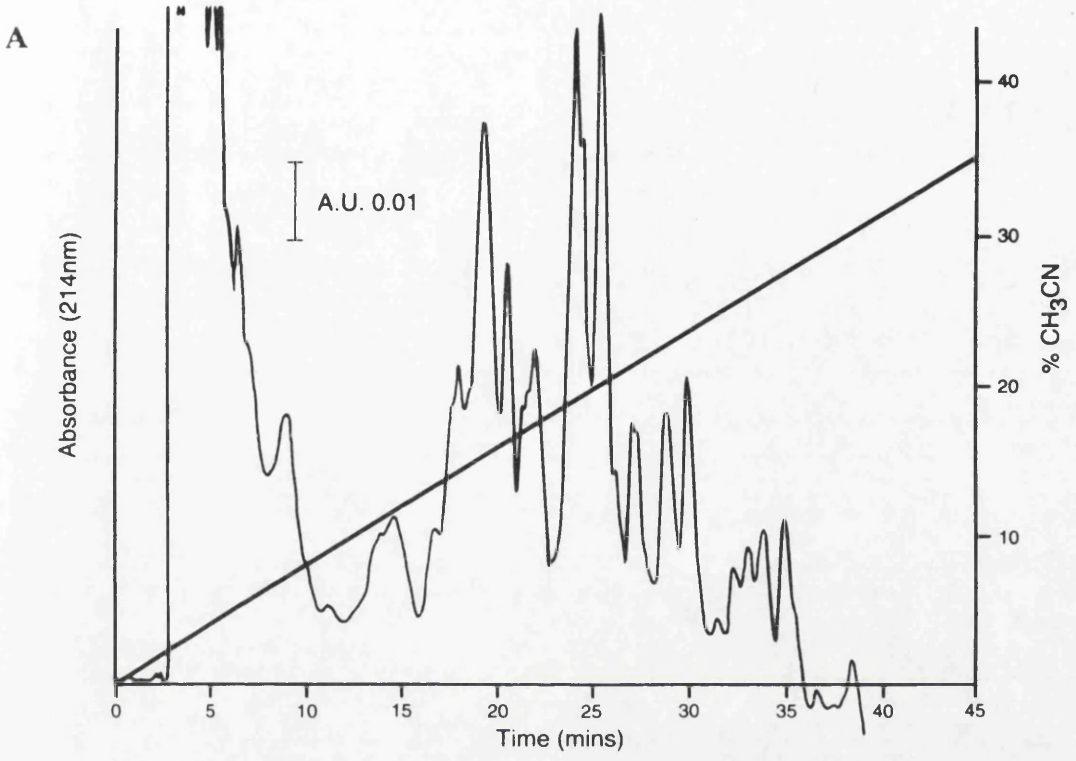
GroEL (~3.5 μM per subunit) was proteolysed in the presence of 2% w/w thermolysin under the conditions described in section 2.12. At time intervals aliquots, containing approximately 5 μg of GroEL, were removed and the reaction stopped by the addition of 2 mM EDTA for 5 minutes. Samples were then analysed on a 15% SDS-PAGE gel (section 2.3) and stained for protein using Coomassie blue as described in section 2.3.1.

Lane 1: markers; BSA, ovalbumin, GAPDH, carbonic anhydrase, trypsinogen, trypsin inhibitor, α–lactalbumin.

Lanes 2-10: GroEL treated with thermolysin for 1, 20, 40, 60, 90, 120, 180, 360 and 1440 minutes respectively.

Fig. 4.6 Reverse phase HPLC profiles, monitored at 214 nm, of thermolysin digested oATP modified GroEL in the absence (A) and presence (B) of ATP

Experimental details are given in section 2.13. Identical quantities of protein were injected (~5 nmoles). The diagonal line in each trace represents the acetonitrile gradient and detection was carried out at 214 nm to detect peptide bonds. AU represents absorbance units.


A
A.U. 0.01
Absorbance (214nm)
% CH_3CN
Time (mins)
0
5
10
15
20
25
30
35
40
45
10
20
30
40

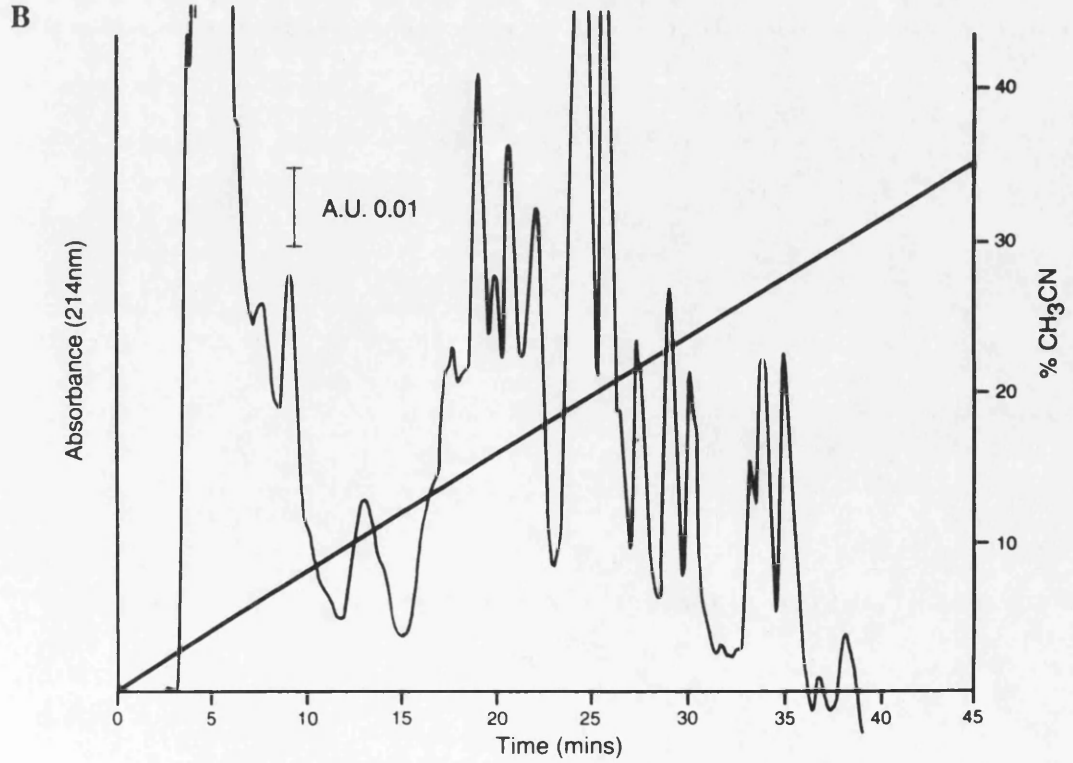
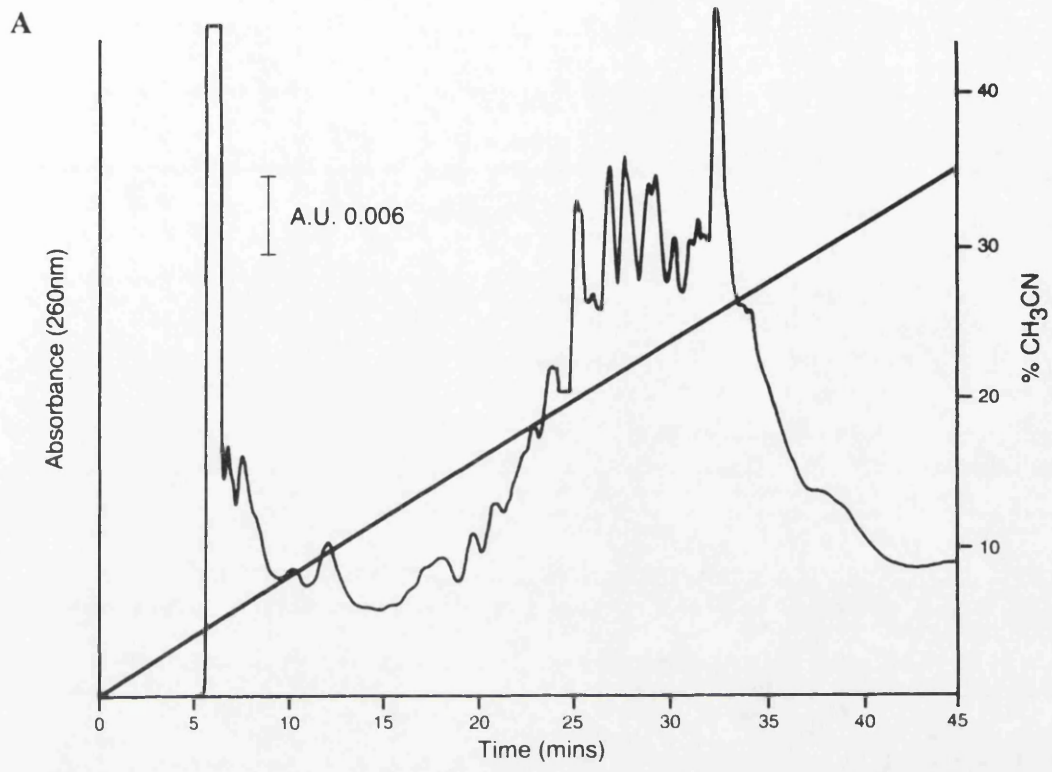

B
A.U. 0.01
Absorbance (214nm)
% CH_3CN
Time (mins)
0
5
10
15
20
25
30
35
40
45
10
20
30
40

Fig. 4.7 Reverse phase HPLC profiles, monitored at 260 nm, of thermolysin digested oATP modified-GroEL in the absence (A) and presence (B) of ATP

Experimental details are given in section 2.13. Identical quantities of protein were injected (~5 nmoles). The diagonal line in each trace represents the acetonitrile gradient and detection was carried out at 260 nm to detect adenine-containing peptides. AU represents absorbance units.


A
A.U. 0.006
Absorbance (260nm)
% CH_3CN
Time (mins)
0
5
10
15
20
25
30
35
40
45
10
20
30
40

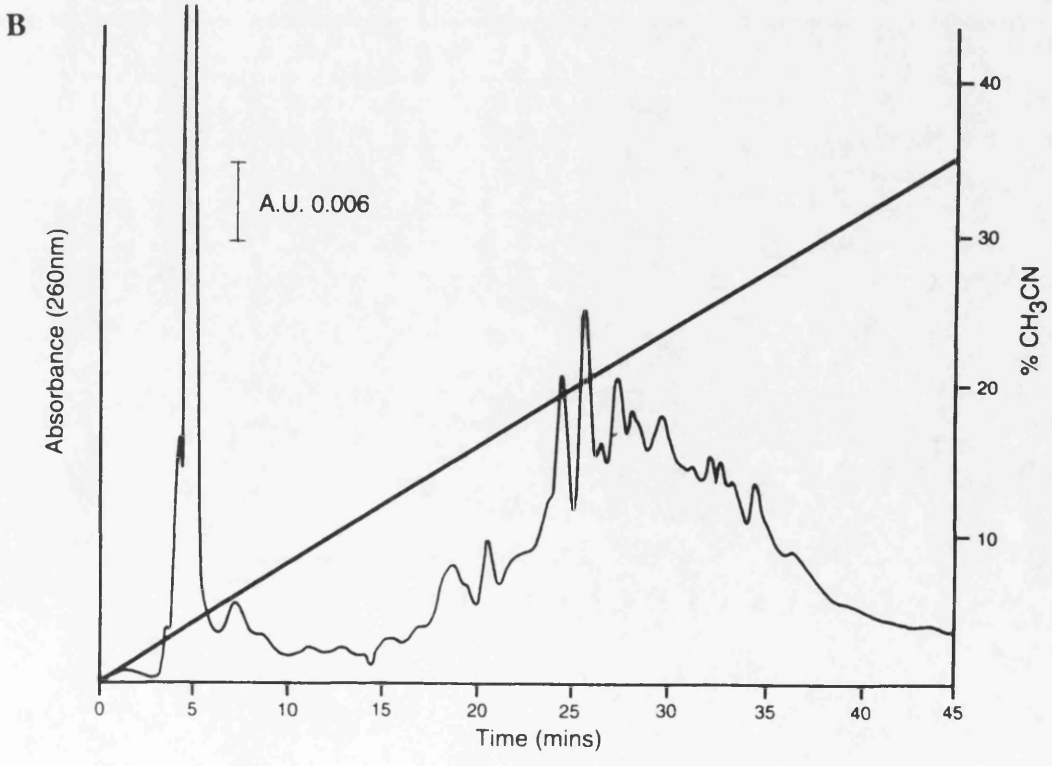

B
A.U. 0.006
Absorbance (260nm)
% CH_3CN
Time (mins)
0
5
10
15
20
25
30
35
40
45
10
20
30
40

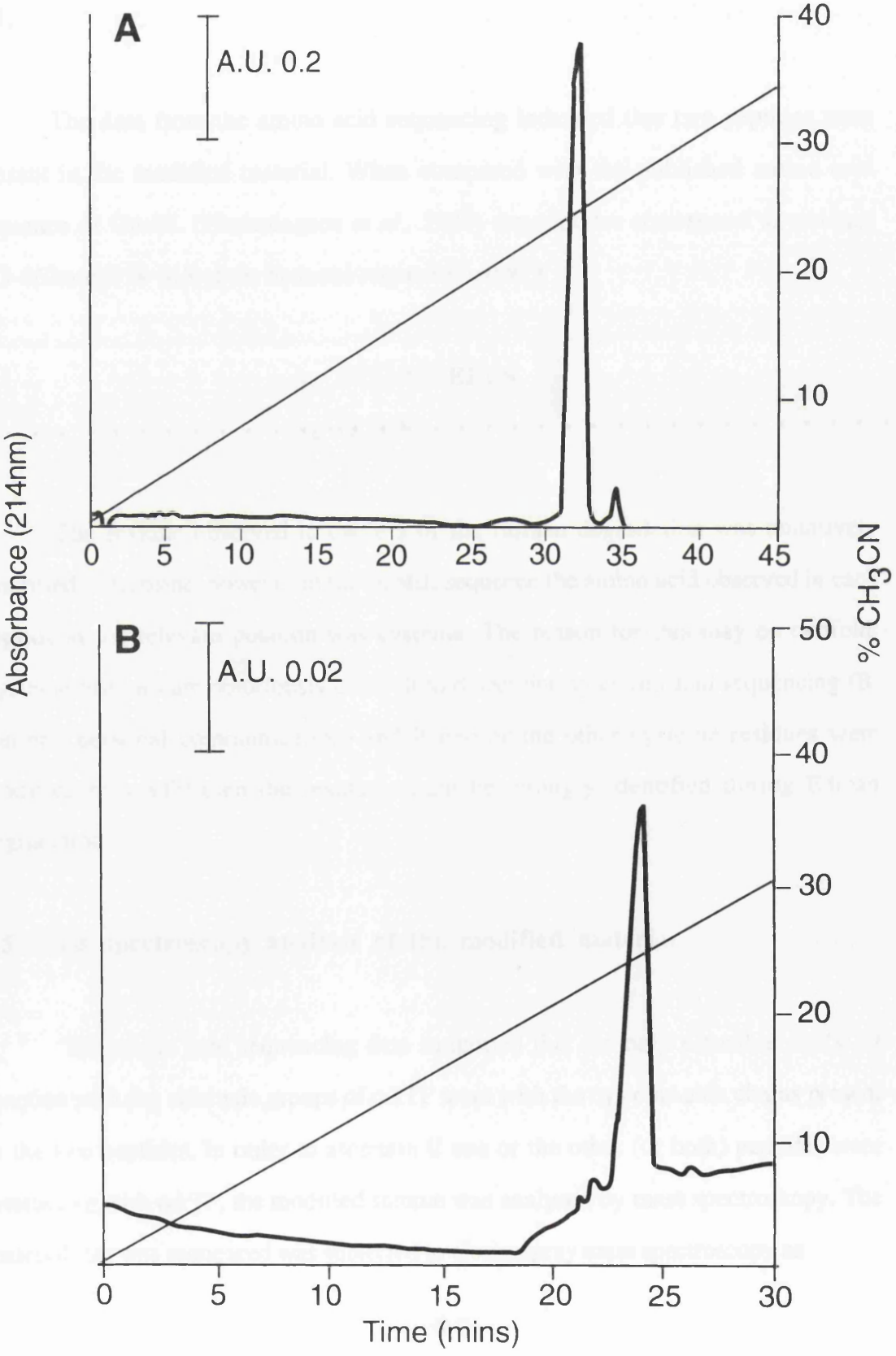
Fig. 4.8 Reverse phase HPLC profiles, monitored at 214 nm, of the modified GroEL peptide

A. Re-chromatography on μBondapak C18 column

The modified peptide eluting after approximately 32 minutes (Fig. 4.7A) was collected and re-chromatographed under the conditions described previously (section 2.13.4) and detection monitored by absorbance at 214 nm.

B Chromatography on a microbore C8 column

The re-chromatographed peptide from Fig. 4.8A was collected and further fractionated on a Pierce microbore C8 column. Elution was carried out with a shallow acetonitrile gradient (section 2.13.4) and detection monitored by absorbance at 214 nm.


A
A.U. 0.2
40
30
20
10
0
5
10
15
20
25
30
35
40
45
B
A.U. 0.02
50
40
30
20
10
0
5
10
15
20
25
30
Time (mins)
Absorbance (214nm)
% CH3CN

4.4.5 Amino acid sequencing of the modified material

The material eluting after approximately 24 minutes from the C8 reverse phase HPLC column (Fig. 4.8B) was collected and subjected to Edman degradation (Table 4.4).

The data from the amino acid sequencing indicated that two peptides were present in the modified material. When compared with the published amino acid sequence of GroEL (Hemmingsen *et al.*, 1988) the peptides correspond to residues 453-462 and 514-518 of the deduced sequence, namely

$_{453}$ **IVLNCGEEPS**

$_{514}$ **ITTEC**

The residue observed in cycle 5 of the Edman degradation was tentatively identified as tyrosine, however in the GroEL sequence the amino acid observed in each peptide at the relevant position was cysteine. The reason for this may be twofold. Cysteine residues are notoriously difficult to detect during amino acid sequencing (B. Dunbar, personal communication) and if one or the other cysteine residues were modified by oATP then the residue would be wrongly identified during Edman degradation.

4.5 Mass spectroscopy analysis of the modified material

The amino acid sequencing data suggested that the only plausible site(s) of reaction with the aldehyde groups of oATP were with the cysteine side chains present in the two peptides. In order to ascertain if one or the other (or both) peptides were interacting with oATP, the modified sample was analysed by mass spectroscopy. The material that was sequenced was subjected to electrospray mass spectroscopy as

Table 4.4

Automated Edman degradation of the oATP modified material

Cycle number	Residue	Yield (pmol)
1	Ala, Val, Ile, Leu	3.7, 11, 28, 15
2	Thr, Val	7.9, 13
3	Thr, Leu	6.2, 10
4	Asn, Glu	9.7, 9.0
5	Tyr	3.5
6	Gly	6.0
7	Glu	5.5
8	Glu	6.0
9	Pro	4.5
10	Ser	2.0

described in section 2.13.6. The material was analysed in both in the positive and negative ionisation mode, in view of the possibility of the peptide containing at least one and possibly three phosphate groups.

The spectrum obtained in the positive ion mode showed a major component with a strong doubly charged ion at m/z of 812.6 which corresponds to a molecular weight of 1623.2. Ions at m/z 823.5 and 831.6 corresponding to the sodium and potassium adducts were also observed (Fig. 4.9A).

The spectrum obtained in the negative ion mode was identical (Fig.4.9B). The species present was a strong doubly charged ion at m/z of 810.6 which again gives a molecular weight of 1623.2. Although other minor peaks were present in both spectra, nothing else of significance was observed.

The molecular mass of 1623 does not represent any of the two peptides with a covalently linked oATP. This molecular mass does however correspond to the mass of the two peptides linked by a disulphide bridge between the cysteine residues.

4.6 Sulphydryl titres of GroEL with 5, 5'-dithiobis(2-nitrobenzoate)

The amino acid sequence of GroEL contains three cysteine residues (Hemmingsen *et al.*, 1988). From the amino acid sequences of the modified peptides the only plausible sites of reaction of oATP with the isolated peptides are with the side chains of the cysteine residues (*i.e.* Cys-457 and / or Cys -518), which would presumably yield hemithioacetal derivatives. These derivatives appear to be stable under conditions used for HPLC and digestion, but are apparently unstable under conditions used for mass spectroscopy analysis and, possibly, Edman degradation.

Fig. 4.9 Electrospray mass spectroscopy of the oATP modified-peptide

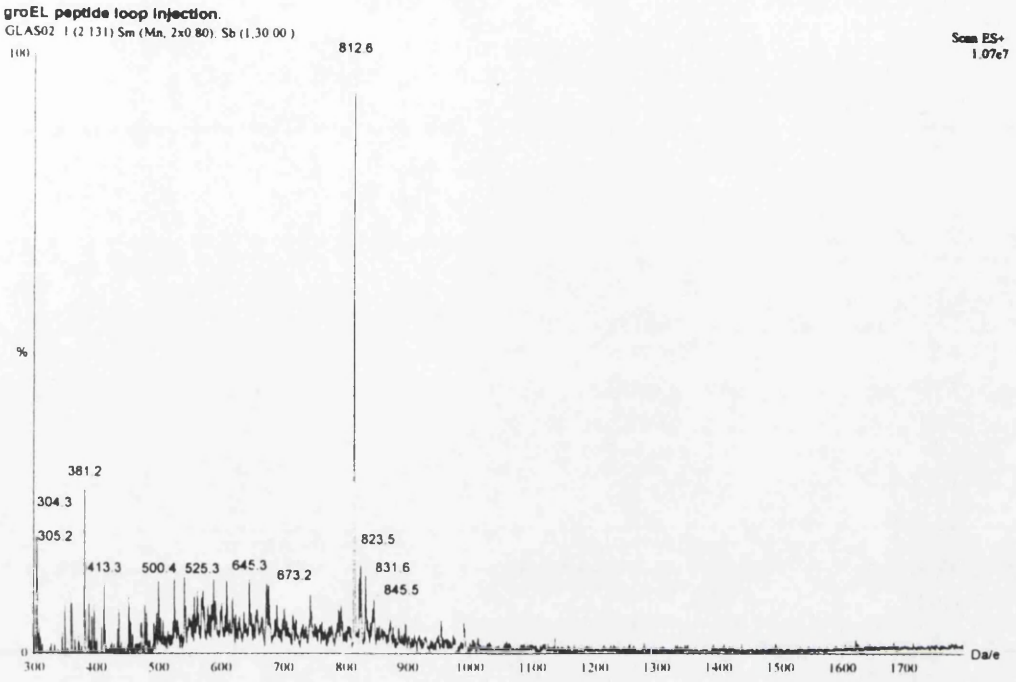
A Electrospray in the positive ion mode

The oATP modified-peptide was collected and further purified by C8 reverse phase chromatography (Fig. 4.8B). The peptide collected was freeze dried and diluted 1:1 with acetonitrile / water before injecting onto mass spectrometer as described in section 2.13.6. The spectrum acquired was in the positive ion mode.

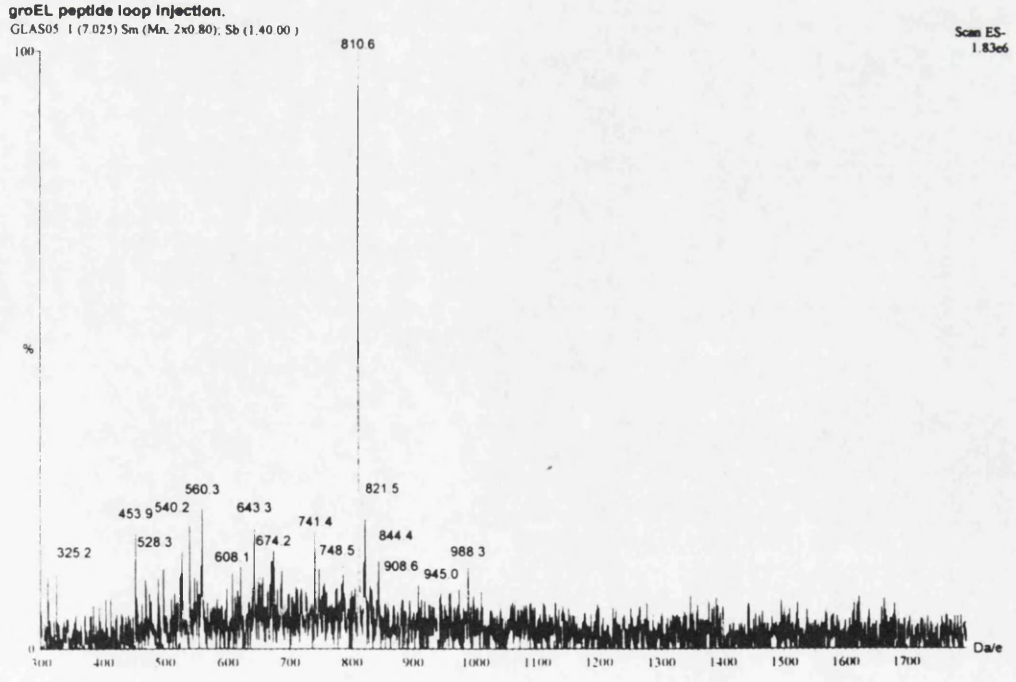
B Electrospray in the negative ion mode

The oATP modified-peptide was collected as described above and analysed in the negative ion mode.

A


groEL peptide loop injection.
GLAS02 1 (2.131) Sm (Mn, 2x0.80); Sb (1,30.00)
Scan ES+
1.07e7
812.6
381.2
304.3
305.2
413.3
500.4
525.3
645.3
673.2
823.5
831.6
845.5
%
Da/e

B


groEL peptide loop injection.
GLAS05 1 (7.025) Sm (Mn, 2x0.80); Sb (1,40.00)
Scan ES-
1.83e6
810.6
560.3
453.9
540.2
643.3
741.4
821.5
325.2
528.3
608.1
674.2
748.5
844.4
908.6
945.0
988.3
%
Da/e

In order to demonstrate an involvement of cysteine side chains in the reaction between GroEL and oATP, GroEL was treated with the thiol specific reagent 5, 5'-dithiobis(2-nitrobenzoate) (DTNB) before and after inactivation with oATP. This assay measures the release of nitrothiobenzoate (NTB) upon reaction of a thiol with DTNB (Fig. 4.10), by the increase in absorbance at 412 nm. Liberation of NTB was calculated using a molar absorption coefficient of 13 600 $M^{-1}cm^{-1}$ (Ellman, 1959). Titratable thiol groups were measured according to section 2.15.

GroEL contains no cysteine residues linked by a disulphide bridge, since all three thiol groups can react with DTNB (Table 4.5). However, the three thiol groups present are free for reaction with DTNB only in the presence of 0.1% SDS suggesting that in the native protein all the cysteine residues are buried within the GroEL structure. When GroEL was reacted with 1 mM-oATP for 6.5 hours, under which conditions > 95% of the ATPase activity was lost, the thiol content was reduced to approximately 1 per chain. ATP protection results in an increased availability of the thiol groups to interact with DTNB. This suggests that GroEL interacts with oATP leading to a loss in thiol groups which is consistent with the results of the sequencing work above.

Table 4.5:- Sulphydryl titration of GroEL

Incubation	Remaining Activity (%)	- SH Modification	
		Native GroEL	+ 0.1% SDS¶
GroEL	100	0.0	3.10
GroEL / ATP	100	0.0	3.20
GroEL / oATP	3	0.0	1.05
GroEL / oATP / ATP	33	0.0	1.94

¶ The results are the average of 3 experiments

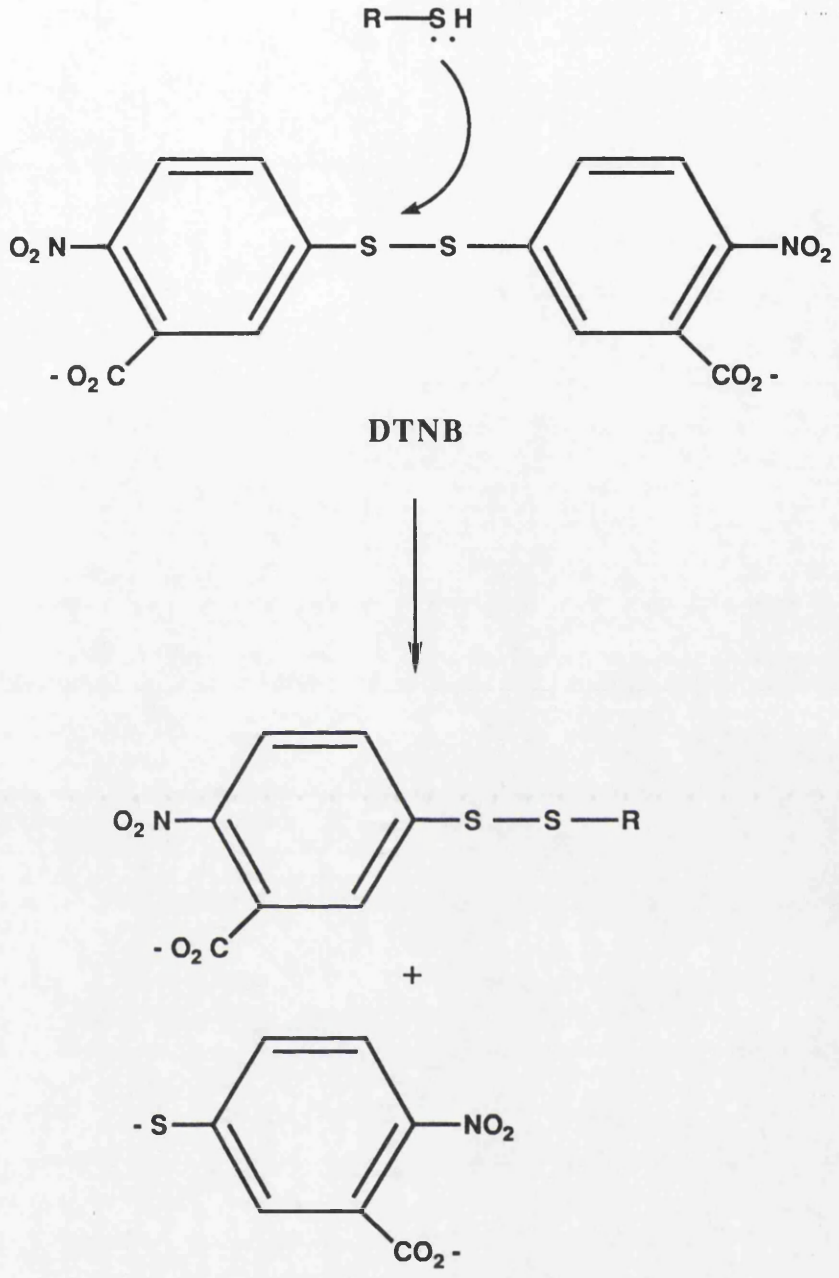


Fig. 4.10 Reaction of thiol with DTNB

4.7 Studies on proteinase K-treated GroEL

Treatment of GroEL with proteinase K leads to the removal of a portion of the C-terminal part of the polypeptide chain, including the highly conserved G/M rich last 13 residues (Langer *et al.*, 1992b; Martin *et al.*, 1993). The truncated form of GroEL retains approximately only 20% of its uncoupled ATPase activity but maintains its characteristic tetradecameric structure and is able to mediate protein folding with rhodanese as a substrate. This truncated form of GroEL was also used in order to study the binding of oATP.

4.7.1 Stoichiometry of oATP incorporation into proteinase K-treated GroEL

GroEL digested in the presence of 2% w/w thermolysin for 60 minutes in 50 mM-TEA, 10 mM-magnesium acetate, pH 8.0 at 25°C (section 2.14.1) was N-terminally intact as shown by amino acid sequencing of a Western blot of the proteolysis (Table 4.6). This is consistent with the work of Langer *et al.* (1992b) who proposed that the proteolysis was from the C-terminus, with the last 16 residues being proteolysed (Martin *et al.*, 1993). The last 16 residues in the GroEL amino acid sequence does not contain any of the 3 cysteine residues known to be present. The extent of incorporation of oATP into proteinase K clipped GroEL (section 2.14.5) was essentially the same as that observed in a parallel experiment with native GroEL (Table 4.7). Incorporation of oATP into proteinase K clipped GroEL was in the order of 2.2 moles of oATP per subunit in inactivated GroEL, with less incorporation observed in ATP protected clipped GroEL. Approximately 1.1 moles of oATP were incorporated into the protected sample. In addition the remaining ATPase activity was protected in the presence of substrate. The ATPase activity in the clipped GroEL is expressed as a percentage of activity remaining over the ATPase activity after treatment with proteinase K. Thus substrate protection could still be demonstrated in the proteolysed form of

GroEL, indicating that ATP binding is unaffected by the absence of these C-terminal residues.

Table 4.6

Automated Edman degradation of proteinase K treated-GroEL

Cycle number	Residue	Yield (pmol)
1	Ala	14
2	Ala	15
3	Lys	15
4	Asp	21
5	Val	13
6	Lys	14
7	Phe	19
8	Gly	9
9	Asn	9.5
10	Asp	10

Table 4.7

Stoichiometry of modification of oATP into proteinase K treated-GroEL

Incubation	Activity (%)	oATP / monomer
Native GroEL / oATP	3	1.76
Native GroEL / oATP / ATP	70	0.60
Clipped GroEL / oATP	13	2.20
Clipped GroEL / oATP / ATP	38	1.1

4.7.2 Stabilisation effects of ATP against proteolysis with proteinase K

The presence of 5 mM-ATP stabilised the GroEL structure towards proteolysis by proteinase K compared to in the absence of ATP (Fig. 4.11A). Although the C-terminal residues were still clipped off, the remainder of the protein was substantially more resistant to further proteolysis. This is likely to be due to an ATP induced conformational change in the GroEL structure which conveys resistance to proteolysis. Interestingly the loss of ATPase activity of GroEL was also protected by the inclusion of ATP (Fig. 4.11B). There was a dramatic loss of ATPase activity upon proteolysis without the prior inclusion of ATP but in its presence, despite an initial dip in activity, approximately 70% of the initial rate of ATP hydrolysis was retained. It appears therefore as though ATP can still bind to GroEL in the absence of the 16 C-terminal residues and be hydrolysed at a rate similar to the native protein when ATP induces a conformational alteration in the structure of GroEL which protects against more extensive proteinase K digestion.

4.7.3 Effect of oATP against proteolysis with proteinase K

The presence of 1 mM-oATP conveyed no protection at all against proteolysis by proteinase K. The stock solution of oATP available did not allow the use of higher concentrations for direct comparison with the protection against proteolysis afforded by 5 mM-ATP, above. Proteolysis appeared to proceed more rapidly in the presence of oATP than in its absence (Fig. 4.12). It is not quite clear why this should occur. One possible explanation is that the subtle structural differences between oATP and ATP may be sufficient to cause slightly different conformational changes in the GroEL structure on interacting with the ATP binding site, thus exposing different regions of the protein towards the protease. The ATPase activity of GroEL in the presence of oATP was not measured in view of the inhibitory effects of the analogue. Different

Fig. 4.11 Proteinase K limited proteolysis of GroEL

A. SDS-PAGE analysis of proteinase K digested GroEL

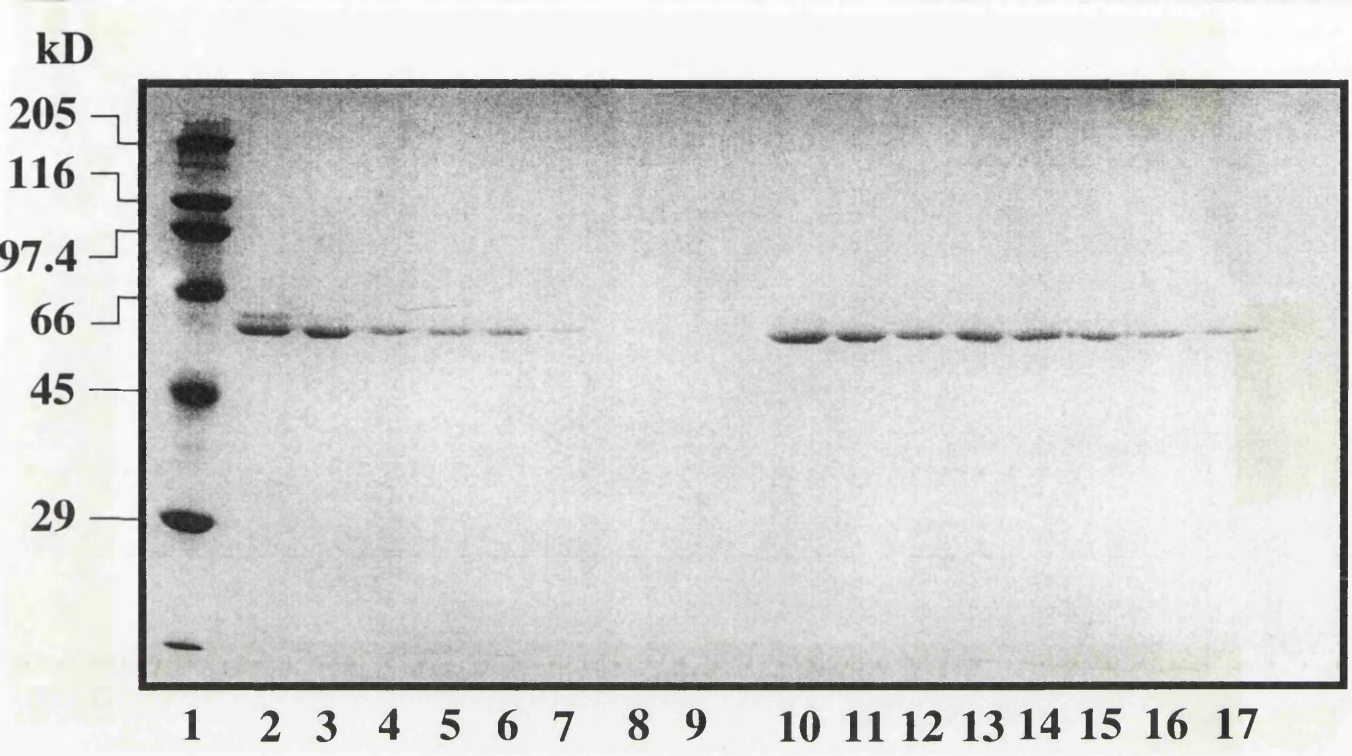
GroEL (~2 μM per subunit) ± 5mM-ATP was proteolysed in the presence of 2% w/w proteinase K under the conditions described in section 2.14.1. At time intervals aliquots, containing approximately 3 μg of GroEL, were removed and the reaction stopped by the addition of 2 mM-PMSF. Samples were then analysed on a 15% SDS-PAGE gel (section 2.3) and stained for protein using Coomassie blue as described in section 2.3.1.

Lane 1: markers (Sigma HMW kit); myosin, β–galactosidase, phosphorylase b, BSA, ovalbumin, carbonic anhydrase.
Lanes 2-9: GroEL treated with proteinase K in the absence of ATP for 5, 30, 60, 90, 120, 180, 240 and 300 minutes respectively.
Lanes 10-17: GroEL treated with proteinase K in the presence of 5 mM-ATP for 5, 30, 60, 90, 120, 180, 240 and 300 minutes respectively.

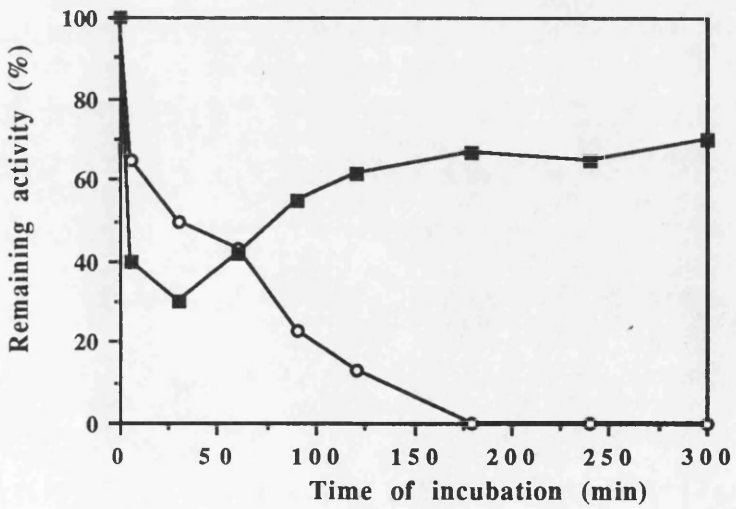
B Effect of limited proteolysis on the ATPase activity of GroEL

GroEL in the absence (O) and presence (■) of 5 mM-ATP was proteolysed in the presence of 2% w/w proteinase K as described above. At time intervals the proteolysis was stopped by the addition of PMSF and aliquots containing ~ 10 μg GroEL were assayed for ATPase activity (section 2.4.1(a)).

A


kD
205
116
97.4
66
45
29
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

B


Remaining activity (%)
100
80
60
40
20
0
0 50 100 150 200 250 300
Time of incubation (min)

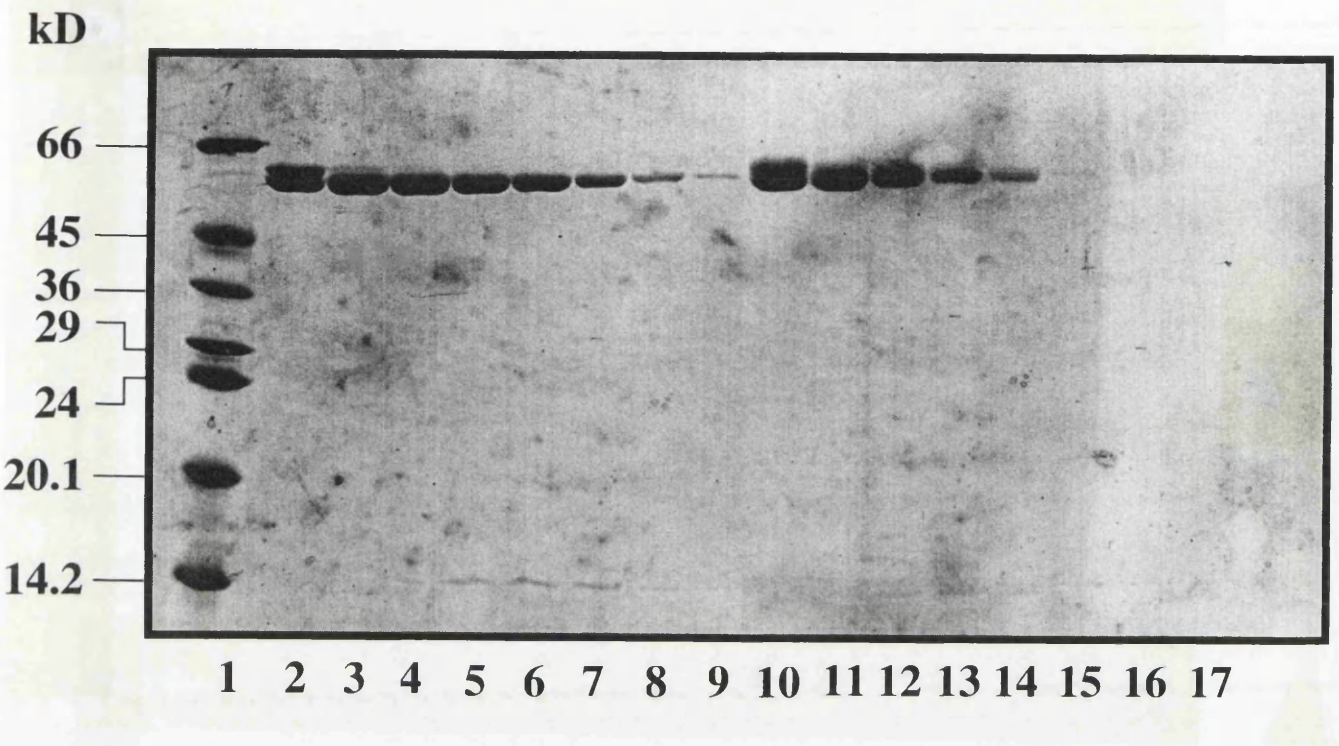


Fig. 4.12 SDS-PAGE analysis of proteinase K treated-GroEL in the presence of oATP

GroEL (~2 μM per subunit) ± 1mM-oATP was proteolysed in the presence of 2% w/w proteinase K under the conditions described in section 2.14.1. At time intervals aliquots, containing approximately 3 μg of GroEL, were removed and the reaction stopped by the addition of 2 mM-PMSF. Samples were then analysed on a 15% SDS-PAGE gel (section 2.3) and stained for protein using Coomassie blue as described in section 2.3.1.

Lane 1: markers; BSA, ovalbumin, GAPDH, carbonic anhydrase, trypsinogen, trypsin inhibitor, α–lactalbumin.

Lanes 2-9: GroEL treated with proteinase K in the absence of oATP for 5, 30, 60, 90, 120, 180, 240 and 300 minutes respectively.

Lanes 10-17: GroEL treated with proteinase K in the presence of 1 mM-oATP for 5, 30, 60, 90, 120, 180, 240 and 300 minutes respectively.

batches of proteinase K were used between the experiments in Figure 4.11 and Figure 4.12 which accounts for the slight differences in rates of proteolysis in the controls.

4.8 Potential active site reagents

In addition to oATP, GroEL was treated with a number of other potential affinity reagents and group specific reagents in order to try to identify the residues involved in ATP binding and hydrolysis. These are discussed below.

4.8.1 5'-Fluorosulphonylbenzoyladenosine

5'-Fluorosulphonylbenzoyladenosine (5'-FSBA) is a synthetic adenosine analogue which has been used as a potential affinity label for a number of adenine binding proteins (Colman *et al.*, 1977). This reagent has a sulphonyl fluoride moiety which can act as an electrophilic agent in covalent modifications of several classes of amino acids including tyrosine, lysine, serine , histidine and cysteine. This broad range of specificity gives a reasonable chance of reaction within any particular nucleotide site.

Treatment of GroEL with 1mM-5'-FSBA under conditions described in section 2.11.2 showed no inhibition of the GroEL ATPase activity even after prolonged incubation for up to 24 hours. This reagent was therefore not suitable for further studies on the GroEL ATP binding site.

4.8.2 4-(iodoacetamido)salicylic acid

Salicylate often competes with ATP so 4-(iodoacetamido)salicylic acid may be considered as a 'pseudo-affinity label', whereby the salicylate moiety directs the reagent towards the ATP binding site. 4-(iodoacetamido)salicylic acid is reactive towards thiol groups and has the advantage of being able to introduce fluorescent groups into

proteins (Price, 1979). There was no noticable loss in GroEL ATPase activity with this reagent (section 2.11.1) in incubations of up to 5 hours, at both 4°C and 25°C.

4.8.3 Pyridoxal 5'-phosphate

Pyridoxal 5'-phosphate has been extensively used to modify lysine residues, which are often found in nucleotide binding sites and especially in phosphate binding loops. Generally, pyridoxal 5'-phosphate inhibits enzymes by the formation of a Schiff base with a reactive lysine. This type of linkage can be readily hydrolysed but irreversible inactivation can be achieved by reduction with sodium borohydride

GroEL (~3.5 μM per subunit) was incubated in 50 mM-TEA, 10 mM-magnesium acetate pH 7.0 with 1mM pyridoxal 5'-phosphate at 25°C followed by fixing with a 10 fold molar excess with sodium borohydride. The reaction was carried out in the dark and at time intervals samples were assayed for ATPase activity. After incubation for up to 3 hours only a small decrease in ATPase activity of 10-15% was observed. No further studies were carried out with this reagent.

4.8.4 Methyl acetimidate

Treatment of GroEL with the lysine specific reagent methyl acetimidate under conditions described in section 4.8.3 showed no inhibition of the ATPase activity with concentrations up to 2 mM even on prolonged exposure.

4.9 Discussion

In order to fully understand the mechanism of action of GroEL several investigators have studied the role of ATP binding and hydrolysis and their relationship to GroEL-mediated protein folding. The binding and hydrolysis of ATP is a highly co-operative process which is enhanced by the presence of the co-chaperonin GroES (Gray and Fersht, 1991; Bochkareva *et al.*, 1992) and is dependent on the presence of both Mg^{2+} and K^{+} for optimum activity (Viitanen *et al.*, 1990). Recent evidence has indicated that the GroEL double-toroid behaves in an asymmetric manner with respect to ATP hydrolysis in the presence of GroES (Todd *et al.*, 1993) and that it is the binding and hydrolysis of the nucleotide (in conjunction with GroES) which drives the 'chaperone machinery' to release polypeptides in a folded or folding-competent state. Despite our (partial) understanding of the chaperoning mechanism, nothing at all is known about the location of the ATP binding site within the GroEL protein. Although the amino acid sequence of GroEL has been elucidated and the quaternary structure solved from electron microscopy the arrangement of the polypeptide chain within the tertiary structure of the protein is unknown. Neither does the published amino acid sequence of the GroEL protein (Hemmingsen *et al.*, 1988) contain the consensus sequence -G-$(X)_4$-G-K-T/S- often found in the phosphate-binding loop of many ATP and GTP binding proteins (Walker *et al.*, 1982; Saraste *et al.*, 1990). The aim of the work described in this chapter was, therefore, to locate and identify the important region(s) of the polypeptide chain involved in the binding of ATP.

The periodate-oxidised ATP analogue oATP differs from ATP only in minimal ribose modification and has been used extensively to specifically react with nucleotide binding sites in proteins. This analogue also appeared to react with the ATP binding site of GroEL and showed properties relevant to an affinity label. The ATPase activity of GroEL was found to be approximately 0.1 sec^{-1} per subunit, in close agreement with other reported values (Viitanen *et al.*, 1990; Todd *et al.*, 1993). As shown in figure

4.2A, the inclusion of oATP with GroEL led to a decrease in activity in a pseudo-first order fashion, which was concentration dependent with respect to oATP. Using 1 mM-oATP the half life of inactivation was approximately 22 minutes with essentially all ATPase activity lost after 2 hours incubation. The demonstration of saturation kinetics provided evidence that oATP forms a dissociable complex with GroEL prior to inactivation. The apparent K_I determined from Fig. 4.2B was 2.1 mM which is comparable to other reported K_I values for oATP inactivation which vary from 6.7 μM for the inactivation of rabbit skeletal muscle phosphorylase kinase (King and Carlson, 1981) to 10 mM for the coupling factor-latent ATPase from *Mycobacterium phlei* (Kumar *et al.*, 1979). In the absence of oATP there was no loss of activity over time and inclusion of ATP led to substantial protection against inactivation (Fig. 4.3). Up to 70% of the ATPase activity could be retained by the inclusion of 5 mM-ATP. This indicates that inactivation is likely to be due to interactions at or near the ATPase active site. In preliminary experiments it was found that the addition of freshly thawed oATP solution to GroEL resulted in a lag period of approximately 2 hours before any substantial inactivation occurred. This lag could be eliminated by preincubation of the oATP at 25°C for 2 hours before addition. However preincubation of GroEL at 25°C before the addition of freshly thawed oATP did not abolish the lag. The reason for this discrepancy is not clear.

Further evidence for the inactivation being at or near the ATPase active site came from the measurement of the stoichiometry of modification. The complete inactivation of GroEL was due to the incorporation of approximately 1.3 moles of oATP per GroEL subunit. This is in line with the proposal that the GroEL tetradecamer has 14 ATP binding sites (Bochkareva *et al.*, 1992). Inclusion of ATP led to a reduction in the incorporation of oATP binding to approximately 0.6 moles per GroEL subunit (Table 4.3). Thus the inactivation of the ATPase activity of GroEL is due to the incorporation of about 1 oATP per GroEL subunit, which can be protected by ATP, although additional reactions appear to occur at other sites. Interestingly, similar levels

of incorporation of oATP into proteinase K treated GroEL (in the presence and absence of ATP) were also observed (Table 4.7). This protein is N-terminally intact but has lost its C-terminus last 16 residues through proteolysis (Martin *et al.*, 1993). The results presented here, however, would suggest that the truncated GroEL can bind ATP in a similar fashion to the full-length protein and may explain why it has similar functional properties both *in vitro*, to refold rhodanese in a GroES and ATP dependent manner, (Langer *et al.*, 1992b), and can also substitute for full-length GroEL *in vivo* (McLennan *et al.*, 1993).

Attempts to locate the site of reaction by a differential peptide-mapping approach proved to be more difficult than anticipated. A modified adenine-containing peptide was located by absorbance at 260 nm; this could be protected by inclusion of ATP (Fig. 4.7). This modified peptide, isolated after reverse-phase chromatography of a complete thermolytic digest of GroEL, was not homogeneous and contained more than 1 peptide sequence. Attempts to separate the peptide species were unsuccessful. The usual site of modification of oATP has been proposed to be with a lysine residue either to form a Schiff base, which can be stabilised by reduction with sodium borohydride to form an irreversible imine linkage, or a dihydroxymorpholino derivative (Easterbrook-Smith *et al.*, 1976; Colman, 1983; Berknov *et al.*, 1990; Rabinkov and Amontov, 1990). Neither of the 2 peptides isolated, corresponding to residues 453-462 (IVLNCGEEPS) and 514-518 (ITTEC) in the GroEL sequence, contained a lysine residue or an arginine residue which is also a potential reactive site. From the sequences of the peptides the only plausible site(s) of reaction are with one or the other cysteine residues present in each peptide. This residue was not identified correctly in either case during peptide sequencing (although cysteines are very difficult to detect during Edman degradation), unlike the other amino acid residues. Indeed the sulphydryl group of cysteine has been proposed to be a potential site of reaction for oATP whereby the product of the reaction is a hemithioacetal (King and Carlson, 1981). On at least one occasion this has been directly demonstrated. The inhibition of avian myeloblastosis virus reverse transcriptase

by oATP is believed to be due to the reactivity of the oATP dialdehyde groups towards essential sulphydryl residues of the protein (Srivastava *et al.*, 1983). Interestingly these authors observed that sodium borohydride treatment reversed the inhibition of oATP. This was not observed upon oATP inactivation of the GroEL ATPase activity. The inactivation of the ATPase activity did not appear to be affected whether or not the protein was treated with sodium borohydride, as inactivation curves were virtually identical (data not shown). Direct evidence for the involvement of cysteine side-chains came from the sulphydryl titration experiments using DTNB. Inactivation of GroEL by oATP led to the blocking of 2 of the 3 thiol groups present in GroEL which were then unable to react with DTNB (Table 4.5). Substrate protection of GroEL resulted in an increased availability of thiol groups for reaction with the sulphydryl-specific reagent, consistent with protection against complete inactivation of the ATPase activity. The decrease in the number of available thiol groups observed upon oATP treatment was not observed upon inclusion of ATP alone, as has been observed previously (Mendoza and Horowitz, 1992), suggesting that this is a property specific to the ATP analogue. Thus from the evidence from the modified peptide sequence(s) and sulphydryl modification it would appear that at least one, and possibly both, of the cysteine residues, Cys 457 and Cys 518, identified in these studies are present in the region of the GroEL ATP binding site.

From the available data, however, it is not possible to predict with any certainty which of the isolated cysteine residues are involved in this function. The reaction of the aldehyde group(s) of oATP at either Cys 457 or Cys 518 (or a cross-link of the oATP with the 2 cysteine residues) to yield hemithioacetal derivatives (Fig. 4. 13) is apparently stable under the conditions used for digestion, HPLC and DTNB treatment but is apparently unstable for mass spectroscopy analysis and, possibly, Edman

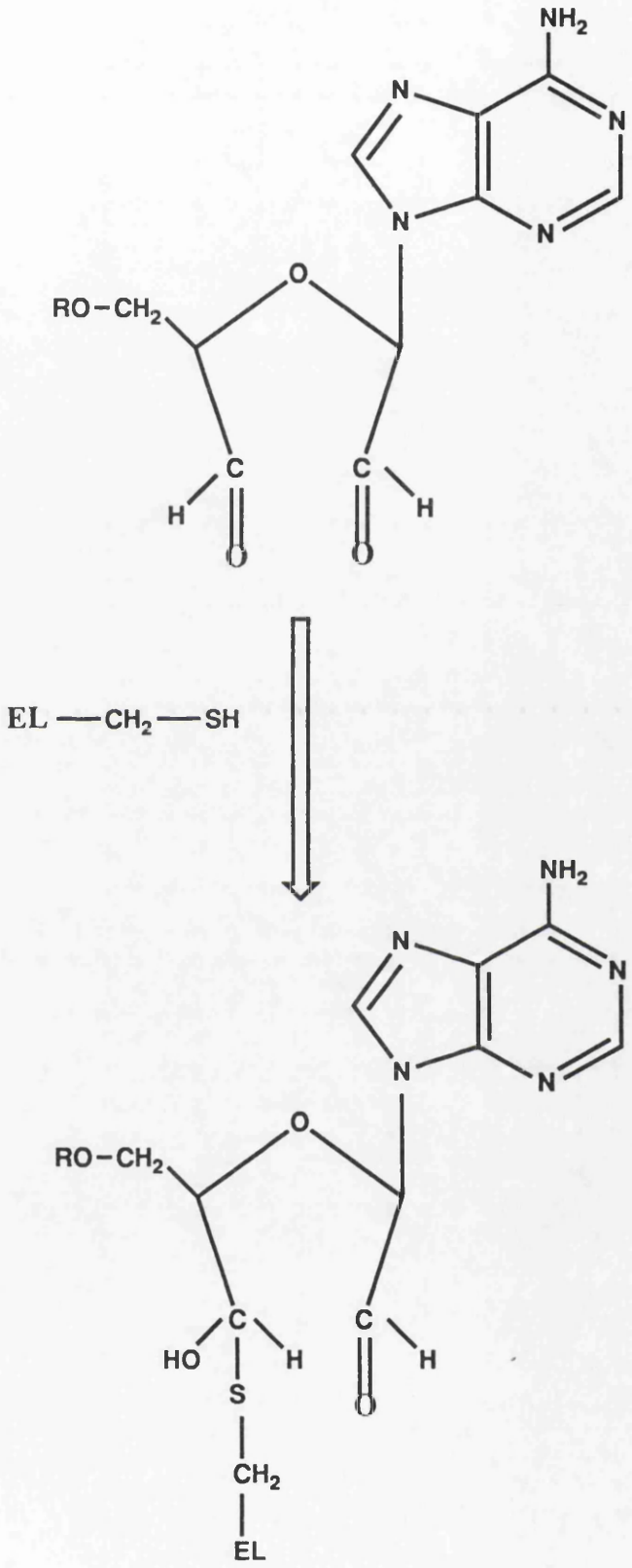


Fig. 4.13 Proposed reaction of oATP with active site cysteine

degradation. Under these conditions it would appear that the adenine moiety is lost and that the presence of the 2 cysteine containing peptides could lead to the formation of a disulphide bridge. There was no evidence for the presence of additional peptides in the correct M_r range for an oATP modified peptide during mass spectroscopy analysis (Fig. 4.9). Interestingly, however, although considerable kinetic evidence exists to suggest that oATP can act as an affinity label for a number of ATP binding proteins, attempts to isolate and identify oATP modified peptides or amino acids have generally proved unsuccessful. In several cases proposed morpholine-like derivatives appear to be unstable during attempts at identification (King and Carlson, 1981; King and Colman, 1983; Berkinov *et al.*,1990). Thus it is feasible that oATP has limited use for the identification of amino acids which form complexes other than reducible Schiff bases with oATP. Clearly a more detailed investigation of the chemistry of the reaction and the stability of the derivatives is required to clarify this situation.

The cysteine residues of GroEL are not conserved so presumably Cys 457 and Cys 518 are not essential for the ATPase catalytic function of the protein; however they are nucleophilic groups which can react with oATP. The indication that Cys 457 and / or Cys 518 may form part of the ATP binding site is consistent with a number of observations. GroEL that has been covalently modified by either of the sulphydryl-specific reagents N-ethylmaleimide or 2-(4'-(iodoacetamido)anilino) naphthalene-6-sulphonic acid is unable to support the GroES and ATP dependent mediated refolding of rhodanese, despite being able to form a binary complex with the unfolded polypeptide (Mendoza and Horowitz, 1992). This may indicate that the ATP-mediated release of the polypeptide from GroEL is affected by the blocking of important cysteine residues present at the ATP binding site. Recently, Bochkareva *et al.* (1994) have demonstrated that the ATPase activity of GroEL can be completely inhibited by another sulphydryl-specific reagent 2, 2'-dipyridyl disulphide and have shown that the nonhydrolysable ATP analogue ^{35}S-ATP-γ–S can specifically form cross-links to Cys 137 in GroEL. Analysis of the stability of single and double mutants in which cysteine

residues are replaced by serine has shown that Cys 518 (identified in these studies) is in close proximity to Cys 137 (Bochkareva *et al.* unpublished results). Using a different affinity labelling approach Martin *et al.* (1993b) have identified a conserved residue, Tyr 477, which can form cross-links to $8N_3$-ATP. These authors have assigned this tyrosine residue to be located in a 40-kD ATP binding domain located in the C-terminal region of the chaperone protein. This proposed ATP binding domain contains both Cys 457 and Cys 518 identified in these studies but does not contain the residue Cys 137 putatively identified by Bochkareva *et al.* (1994) to be responsible for binding the γ–phosphate group of ATP (although this bond may have been hydrolysed in the work of Martin *et al.* (1993b)). Until the three dimensional structure of the protein is solved one can only speculate on the location of the ATP binding site. Alignments of the amino acid sequences of members of the chaperonin 60 family (including the eukaryotic cytosol TRiC class) show 2 regions of especially high homology (Lewis *et al.*, 1992). A conserved feature of these proteins is ATP binding and hydrolysis, so comparisons of sequences should help to identify regions involved in this function. The first block of homology between amino acids 80-100 show a highly conserved motif centred around -GDGTT(T/S)-. Lewis *et al.* (1992) have identified this region as a possible β–phosphate ATP binding domain closely related to that found in cAMP dependent protein kinase. This region is, however, outwith the putative 40-kD ATP binding domain identified by Martin *et al.* (1993b), who themselves have proposed a region spanning residues 164-172 and 241-251 of the GroEL sequence to be responsible for binding of triphosphate groups. Sequence alignments of these regions of GroEL show close homology to other known chaperonin 60 proteins and contain a conserved potential active site lysine residue located at position 167 in GroEL (Fig. 4.14). Site-directed mutagenesis of this residue may provide important clues regarding its location relative to the ATP binding site. These sequences, however, are unrelated to members of the TRiC sub-class of chaperonins. The second block of homology identified by Lewis *et al.* (1992), present in both the classical chaperonin 60s and the TRiC class, is

	164 * 172
GroEL (*E. coli*)	**AMDKVGKEGVIT**
RuBP (*T. aestivum*)	**A IDKVGPDGVLS**
Hsp60 (*S. cerevisiae*)	**AMEKVGKEGVIT**
Hsp60 (*H. sapien*)	**AMEKVGKEGVIT**

	241 251
GroEL (*E. coli*)	**KAGKPILIIAE**
RuBP (*T. aestivum*)	**QLRCPLFIVAE**
Hsp60 (*S. cerevisiae*)	**QSRRPLLIIAE**
Hsp60 (*H. sapien*)	**QSRRPILI IAE**

Fig. 4.14 Sequence alignments of proposed ATP triphosphate binding regions in chaperonin 60 proteins.

These putative ATP binding regions of GroEL were proposed by Martin *et al.* (1993b). The numbering is according to the *E. coli* protein. *Bold faces* are residues conserved in all four sequences. * indicates a highly conserved Lysine residue which may be present in an ATP phosphate binding loop. These alignments are taken from the sequences reported by Lewis *et al.* (1992) and Tsugeki *et al.* (1992).

between residues 400-500, where there are no significant insertions or deletions, which contain a conserved -V(A/P)GGG- motif. This region may also play an important role in binding parts of the ATP molecule. The definitive answer to this question will come from X-ray crystallography studies and site-directed mutagenesis of potential active site amino acids.

In conclusion, a region of the GroEL protein has been identified which may be responsible for ATP binding. This region is located towards the C-terminus of the protein and is dependent on the presence of Cys 417 and / or Cys 518. It should be noted that the reactive functional groups of oATP are in the ribose moiety of the adenine nucleotide, whereas Bochkareva *et al.* (1994) and Martin *et al.* (1993b), using ATP analogues with different reactive groups, identified ATP binding regions in the GroEL molecule towards the N and C-terminus, respectively. These results, taken together with the conserved amino acid regions, indicate that the tertiary structure of each subunit must be as such to bring these regions of the polypeptide chain into close proximity so as to create an ATP binding site. One possible arrangement is that amino acid residues towards the N-terminal half of GroEL are involved in the binding of phosphate groups of ATP, as has been observed for a number of ATP binding proteins (Walker *et al.*, 1982; Saraste *et al.*, 1990), whilst the amino acids near the C-terminus are involved in the binding of the sugar moiety, and the purine ring. A similar type of structural arrangement exists in several ATP binding proteins including sugar kinases, actin and Hsp70 proteins (Bork *et al.*, 1992). It remains a task for future work to fully characterise the amino acids involved in the ATP binding site of GroEL.

As an additional point of interest, the reported specific absorption coefficient values of GroEL at 280 nm (for a 1 mg/ml solution) show large discrepancies, varying between 0.21 - 0.416. Hayer-Hartl and Hartl (1993) proposed that this could be accounted for by the presence of variable amounts of contaminating polypeptides between preparations, although the existence of mutant GroEL polypeptides was not

ruled out. There is obviously a need to determine correctly the concentrations of GroEL solutions in order to study stoichiometry of interactions and understand the mechanism of action of the chaperone protein. The value used in these experiments was the one quoted by Price *et al.* (1991) of 0.285. This would seem to be a good estimate for the GroEL preparations used here since the estimation of thiol groups present in GroEL by DTNB, using this specific absorption coefficient, was in agreement with the number (*i.e.* 3) in the published amino acid sequence (Hemmingsen *et al.*, 1988). Thus DTNB titration may also prove to be a useful tool to help determine the validity of specific absorption coefficients of GroEL preparations.

Chapter 5 The unfolding and attempted refolding of GroEL

5.1 Introduction

GroEL, in conjunction with ATP and its co-chaperonin GroES, is known to be able to assist in the refolding and assembly of a number of chemically denatured polypeptides. The mechanism of assembly of GroEL, and GroEL homologues, themselves also pose interesting problems. It has been proposed that these proteins may undergo a process of 'self-chaperoning' in order to attain native structure (Cheng *et al.*, 1990; Lissin *et al.*, 1990). In order to examine some aspects of these processes, the unfolding of GroEL in the presence of guanidinium chloride (GdnHCl) was studied monitoring changes in ATPase activity and in secondary, tertiary and quaternary structure. In order to initiate the refolding process the denaturant was removed by dialysis which is a widely used technique for studying protein refolding *in vitro*.

5.2 Unfolding of GroEL by GdnHCl

Unfolding of GroEL was studied in the buffer used by Lissin *et al.* (1990), namely 50 mM-potassium phosphate, 20 mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5. Solutions of GroEL were incubated in the presence of GdnHCl for 15 minutes at 20°C before readings were taken; there were no further significant changes upon incubation for a further 45 minutes.

5.2.1 Changes in secondary structure

The changes in GroEL secondary structure brought about by GdnHCl was monitored by circular dichroism in the far-uv region (190-250 nm). This region of the spectrum is dominated by the contributions of the peptide bonds and as such is useful to monitor changes in secondary structure in the course of structural transitions.

The CD spectra of GroEL in the absence of the denaturing agent and in the presence of selected concentrations of GdnHCl are shown in Figure 5.1A. In the absence of GdnHCl, the secondary structure content, derived from the spectrum over the range 240 -190 nm, corresponds to 44 ± 1% α–helix, 29 ± β–sheet, 27 ± 2% remainder. These values were calculated using the method of Provencher and Glöckner (1981) which uses a linear combination of the CD spectra of reference proteins with known secondary structure for a direct analysis of a given protein. The α-helix content is somewhat lower from the value reported by Lissin *et al.* (1990) (57%) who used the same method of secondary structure estimation. At least part of this discrepancy may arise from the difficulties in determining accurate concentrations of solutions of GroEL (Price *et al.*, 1991).

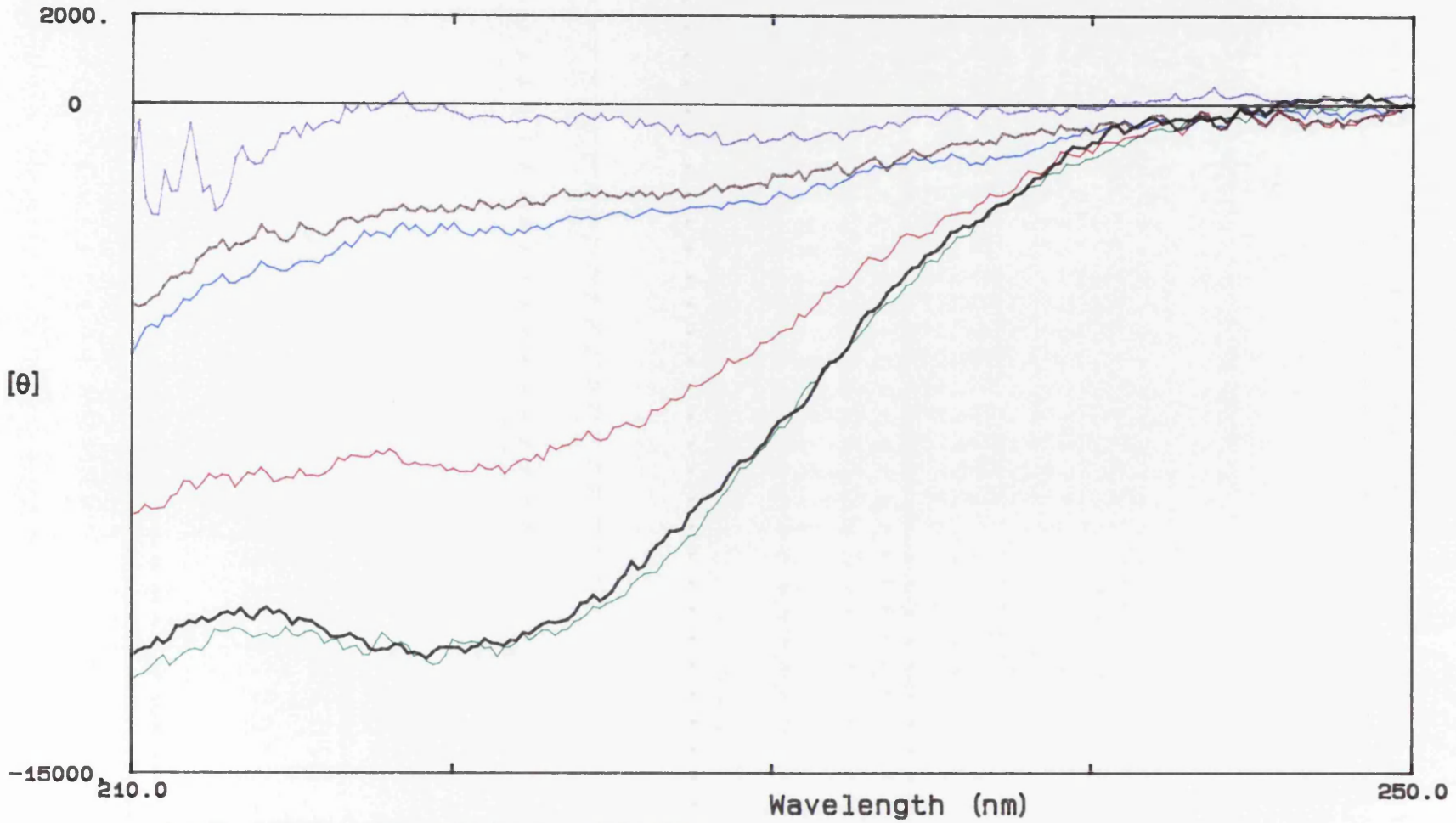
The incubation of GroEL with increasing concentrations of GdnHCl causes a marked loss in secondary structure (Fig 5.1A) whereby a random coil type structure is adopted in the presence of upwards of 2.0 M-GdnHCl. The changes in ellipticity at 225 nm on addition of the denaturant show that there is a sharp transitional change in the region from 1.0 - 1.5 M GdnHCl, with a midpoint in the region of 1.35 M GdnHCl (Fig. 5.1B). In preliminary experiments at 2°C it was observed that the mid-point of transition was lowered to 1.25 M GdnHCl, supporting the observation of Lissin *et al.* (1990) that the chaperone protein is rather less stable at lower temperatures.

5.2.2 Changes in tertiary structure

The tertiary structural changes of GroEL in the presence of the denaturing agent was measured by fluorescence. The fluorescence of proteins originates from the aromatic amino acids, phenylalanine, tyrosine and tryptophan. Fluorescence emission is observed when an excited electron returns from the first excited state to the ground state whereby, due to energy loss, the emission is shifted to a longer wavelength

Fig. 5.1A Changes in the far-uv CD spectrum of GroEL in the presence of GdnHCl

Spectra were obtained in a buffer of 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 at 20°C. Unfolding of GroEL (50 μg / ml) in the presence of varying concentrations of GdnHCl was carried out as described in section 2.8.2. Spectra refer to GdnHCl concentrations of 0, 0.5, 1.3, 1.5, 2.0 and 6.0 M, respectively. Spectra were recorded in a pathlength of 1 mm (section 2.9.1) by Sharon Kelly in the laboratory of Dr. Nicholas Price at Stirling University.


2000.
0
[θ]
-15000
210.0
250.0
Wavelength (nm)
GroEL 0MGDN 0.05cm 10/10/91 CORROPT
GroEL 0.5MGDN 0.05cm 10/10/91 CORROPT
GroEL 1.3MGDN 0.05cm 10/10/91 CORROPT
GroEL 1.5MGDN 0.05cm 10/10/91 CORROPT
GroEL 2MGDN 0.05cm 10/10/91 CORROPT
GoEL 6MGDN 0.05cm 10/10/91 CORROPT

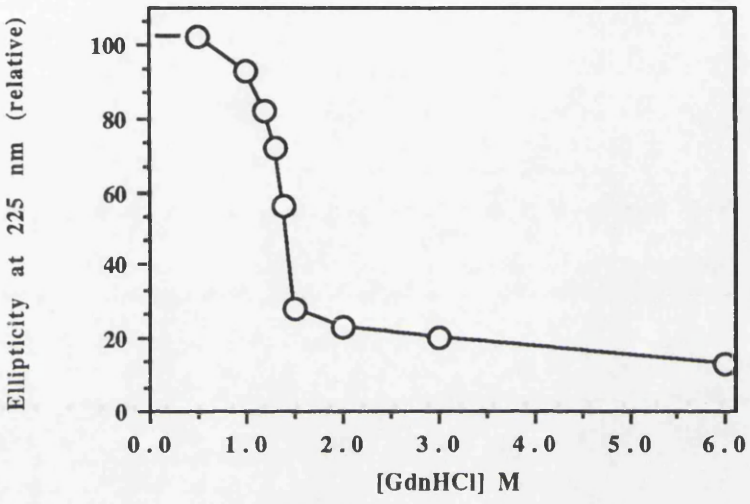


Fig. 5.1B Changes in ellipticity at 225 nm in the far-uv CD spectra of GroEL in the presence of GdnHCl

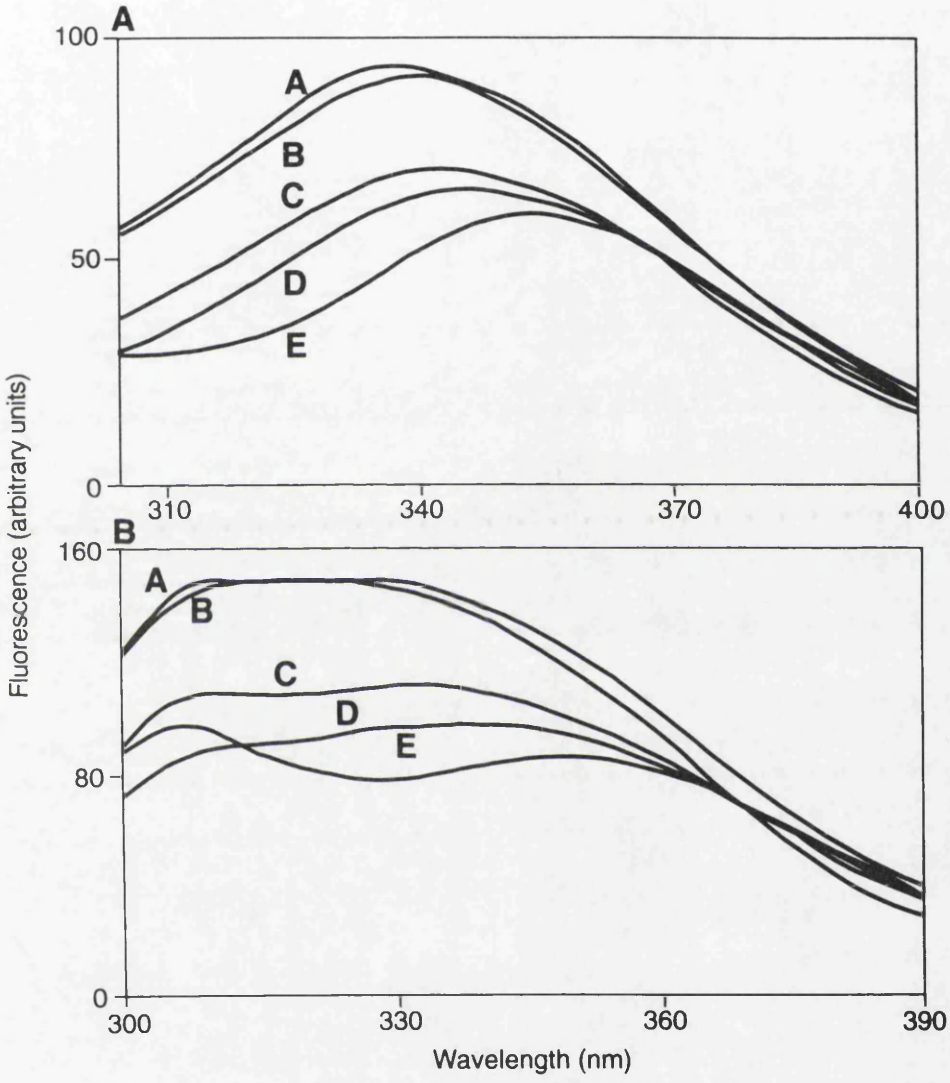
The changes of ellipticity at 225 nm of the CD spectra of GroEL in varying concentrations of GdnHCl (Fig. 5.1A) are shown relative to that of GroEL in the absence of GdnHCl.

compared to the absorption of the chromophore. The contribution of phenylalanine is generally very small compared to tyrosine and especially tryptophan, which is the dominant amino acid. This technique permits monitoring of the environment in which these amino acids exist. The aromatic amino acid content of GroEL is disputed. The published amino acid sequence of GroEL, derived from its cDNA sequence, contains 7 tyrosine residues and no tryptophan residues (Hemmingsen *et al.*, 1988). Price *et al.* (1991) have proposed that 1 tryptophan residue is present per GroEL subunit with 7 tyrosine residues, although this may result from contaminating protein(s) (Hayer-Hartl and Hartl, 1993).

The fluorescence emission spectra of GroEL in varying concentrations of GdnHCl were determined with exitation at both 290 nm (Fig. 5.2A) and 280 nm (Fig. 5.2B). In each case it is clear that major changes occur in the spectrum over the range 1.0 M - 1.5 M GdnHCl . There is a shift in wavelength maximum as the protein unfolds. This is easier to appreciate in figure 5.2A where tryptophan fluorescence is selectively excited, than in figure 5.2B where both tyrosine and tryptophan contribute to the emission. In the absence of GdnHCl, the wavelength of maximum emission is 335 nm; this is shifted to 355 nm, the value characteristic of solvent exposed tryptophan, in 6 M GdnHCl. The unfolding of GroEL over the narrow range (1.0 -1.5 M) of GdnHCl is demonstrated in figure 5.2C which shows the changes in intensity of fluorescence at 315 nm for excitation at both 280 nm and 290 nm.

Fig. 5.2 Changes in the fluorescence of GroEL in the presence of GdnHCl

Spectra were obtained in a buffer of 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 at 20°C. Unfolding of GroEL (50 μg / ml) in the presence of varying concentrations of GdnHCl was carried out as described in section 2.8.2. Spectra A, B, C, D and E refer to GdnHCl concentrations of 0, 1.0, 1.3, 1.5 and 6.0 M, respectively. **Fig. 5.2A** and **5.2B** refer to excitation at 290 nm and 280 nm, respectively.


A
100
50
0
A
B
C
D
E
310
340
370
400
B
160
80
0
A
B
C
D
E
300
330
360
390
Fluorescence (arbitrary units)
Wavelength (nm)

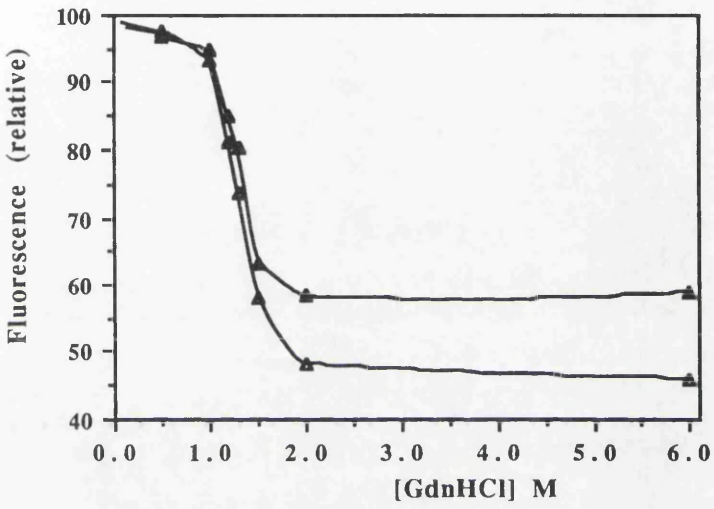


Fig. 5.2C Changes in fluorescence at 315 nm of GroEL in the presence of GdnHCl

Relative changes in fluorescence at 315 nm of GroEL in the presence of GdnHCl when excited at 280 nm (▲) and 290 nm (Δ). See figure legend 5.2 A/B for experimental details.

5.2.3 Changes in quaternary structure

Changes in the quaternary structure of GroEL were measured by light scattering as described in section 2.9.3. Molecular weight ratios of GroEL in the presence of GdnHCl, M_i, to that in the absence of GdnHCl, M_o, were determined by the equation (Parr and Hammes, 1975)

$$\frac{Mi}{Mo} = \frac{(no)^2 (\partial n / \partial co)^2 (c\,o)\,(Ii,90)}{(ni)^2 (\partial n / \partial ci)^2 (c\,i)\,(Io,90)}$$

where n is the refractive index, $\partial n / \partial c$ is the refractive index increment, c is the protein concentration, and I_{90} is the light scattering intensity at 90°.

The changes in GroEL molecular weight at different GdnHCl concentrations was measured by changes in intensity of light scattering at 360 nm. This data shows that a marked loss in quaternary structure took place, again over the range 1.0 - 1.5 M GdnHCl (Fig. 5.3).

5.2.4 Changes in GroEL ATPase activity

The ATPase activity of GroEL in the presence of GdnHCl was measured by a coupled quench assay of the type described by Johnson and Price (1988), in view of the likely effects of the denaturing agent on the coupling enzymes (section 2.4.1(b)). The validity of this method was confirmed by control experiments. The GroEL ATPase activity was lost in the presence of GdnHCl; essentially all activity was lost by 0.75 M GdnHCl (Fig. 5.4).

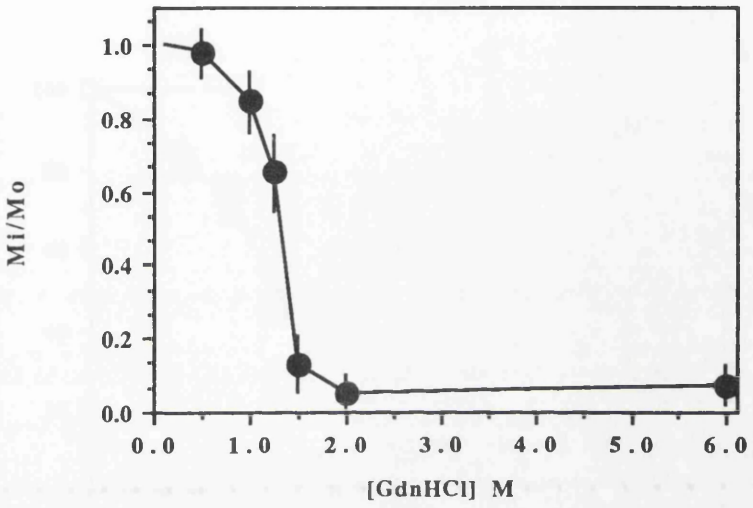


Fig. 5.3 Changes in M_r of GroEL in the presence of GdnHCl

Data were obtained in a buffer of 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 at 20°C. Unfolding of GroEL (50 μg / ml) in the presence of varying concentrations of GdnHCl was carried out as described in section 2.8.2. The changes in M_r were determined by light scattering as described in sections 5.2.3 and 2.9.3. The experimental points represent the means of 4 determinations.

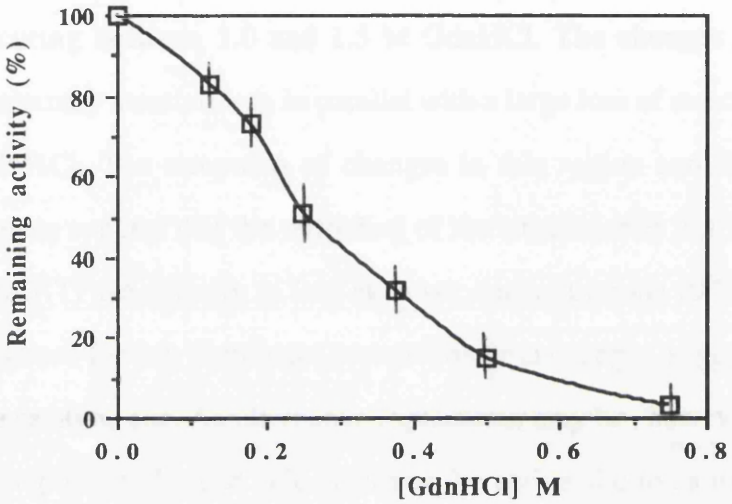


Fig. 5.4 Changes in the ATPase activity of GroEL in the presence of GdnHCl

GroEL was unfolded in varying concentrations of GdnHCl at 20°C in a buffer of 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 as described in section 2.8.2. After treatment with GdnHCl, the ATPase activity of GroEL was measured as described in section 2.4.1 (b).

5.2.5 Comparison of the various techniques

The various techniques used for monitoring the structural changes during the unfolding of GroEL are compared in figure 5.5, where the data are normalised to show the changes as a percentage of the maximum change observed. The various measurements from CD, fluorescence and light scattering all show major changes occuring between 1.0 and 1.5 M GdnHCl. The changes in secondary, tertiary and quaternary structure run in parallel with a large loss of structure over a narrow range of GdnHCl. The steepness of changes in this region and the lack of any discernible plateau suggest that the unfolding of the tetradecamer is a highly cooperative process. The ATPase activity is lost at lower concentrations (0.75 M), where there is little evidence for any significant overall structural changes, suggesting to a greater lability of the catalytic site. An alternative explanation may be , however, that the denaturant binds to important charged side chains at the active site to cause inhibition of the ATPase activity.

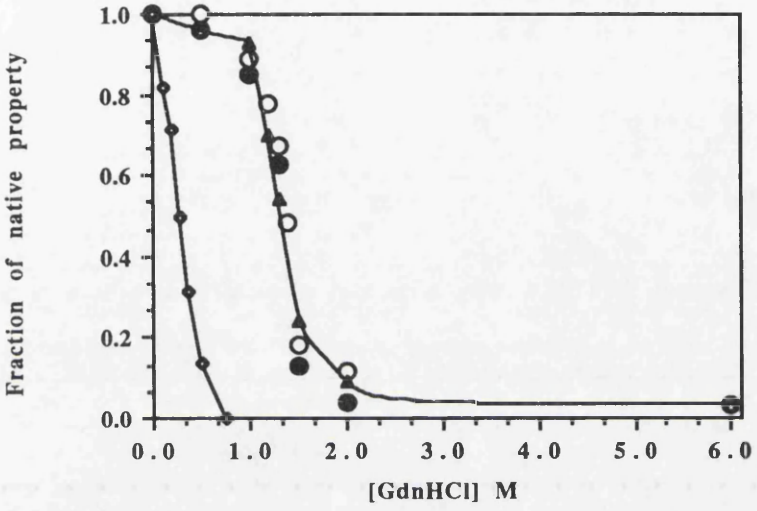


Fig. 5.5 Changes in the ATPase activity and structural parameters of GroEL in the presence of GdnHCl

The data are normalised to show the difference between GroEL in the absence and presence of 6 M GdnHCl as 100%. (◊) changes in ATPase activity, (O) changes in ellipticity at 225 nm, (●) changes in M_r from light scattering, (Δ) changes in fluorescence at 315 nm (excited at 290 nm). The concentration of GroEL was 50 μg / ml for structural measurements and 40 μg / ml for activity measurements. The buffer was 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 at 20°C.

5.3 Refolding of GdnHCl denatured GroEL

The refolding of denatured proteins can be initiated by the removal of the structure perturbing agent , whether by dilution, dialysis or by a pH-jump. Attempted refolding of GroEL which had been denatured in GdnHCl (above) was undertaken by the dialysis technique. Denatured GroEL was dialysed against 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 (section 2.8.3) and changes in structural parameters and ATPase activity were measured as described in the previous section.

5.3.1 Refolding of GroEL after denaturation in low concentrations of GdnHCl

When the initial concentration of GdnHCl was 1.0 M or below there was essentially complete restoration of secondary, tertiary and quaternary structure of GroEL as judged by CD, fluorescence and light scattering, respectively. In addition the ATPase activity of the protein, which was lost by 0.75 M GdnHCl, was restored (Fig. 5.6). However when the initial concentration of GdnHCl was raised progressively above 1.0 M, there was a decreasing extent of regain of ATPase activity (Fig. 5.6).

5.3.2 Refolding of GroEL after denaturation in higher concentrations of GdnHCl

The renaturation of GroEL after incubation with concentrations of GdnHCl above 1.0 M will be discussed in this section.

5.3.3 Secondary structure changes

Irrespective of the initial concentration of GdnHCl, the regain of secondary structure as revealed by CD was almost complete. GroEL which was initially denatured in 6.0 M GdnHCl and had been completely unfolded in a random coil type structure regained approximately 90% of its secondary structure upon dialysis, as judged by ellipticity at 225 nm (Fig. 5.7). It was also found in preliminary experiments that when dialysis was performed at 2°C instead of 20°C the regain of secondary structure was less efficient (57% of the ellipticity at 225 nm was restored).

5.3.4 Tertiary structure changes

The loss in the ability of GroEL to regain its ATPase activity at concentrations above 1.0 M GdnHCl (Fig. 5.6) correlated with a progressive loss of ability of GroEL to regain correct tertiary structure as monitored by the wavelength of maximum emission when excited at 290 nm (Fig. 5.8A). When the initial concentration of GdnHCl was 6.0 M (corresponding to complete unfolding of the protein) the wavelength of maximum emission after dialysis was 345 nm. This corresponds to an intermediate state between the native protein (λ maximum emission 335 nm) and the fully unfolded state (λ maximum emission 355 nm). The fluorescence emission spectra when excited at 280 nm (Fig. 5.8B) also provides evidence for the attainment of an intermediate state when the denaturant was removed.

5.3.5 Quaternary structure changes

The measurements of quaternary structure by light scattering revealed that under the conditions employed the native quaternary structure was not regained, since in most experiments the intensity of light scattering was only some 40 - 50% of a control sample which had been incubated in the absence of GdnHCl and subsequently

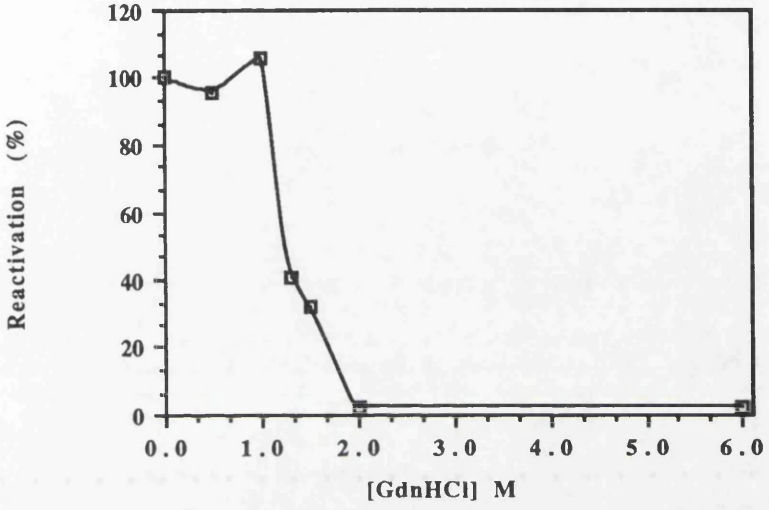


Fig. 5.6 Reactivation of GroEL after incubation in GdnHCl

GroEL (50 μg / ml) was incubated with the stated concentrations of GdnHCl for 15 minutes at 20°C (section 2.8.1); the denaturing agent was then removed by dialysis as described in section 2.8.3. The degree of reactivation is expressed as a regain of ATPase activity (section 2.4.1 (a)), relative to a control sample incubated in the absence of denaturant.

Fig. 5.7 Changes in the far-uv CD spectra of GroEL after removal of GdnHCl by dialysis

Spectra were obtained in a buffer of 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 at 20°C. Unfolding of GroEL (50 μg / ml) in the presence of GdnHCl was carried out as described in section 2.8.2. GdnHCl was then removed by dialysis (section 2.8.3). The initial concentrations of GdnHCl were 0, 0.5, 1.3, 1.5, 2.0 and 6.0 M, respectively. Spectra were recorded in a pathlength of 1 mm (section 2.9.1) by Sharon Kelly in the laboratory of Dr. Nicholas Price at Stirling University.

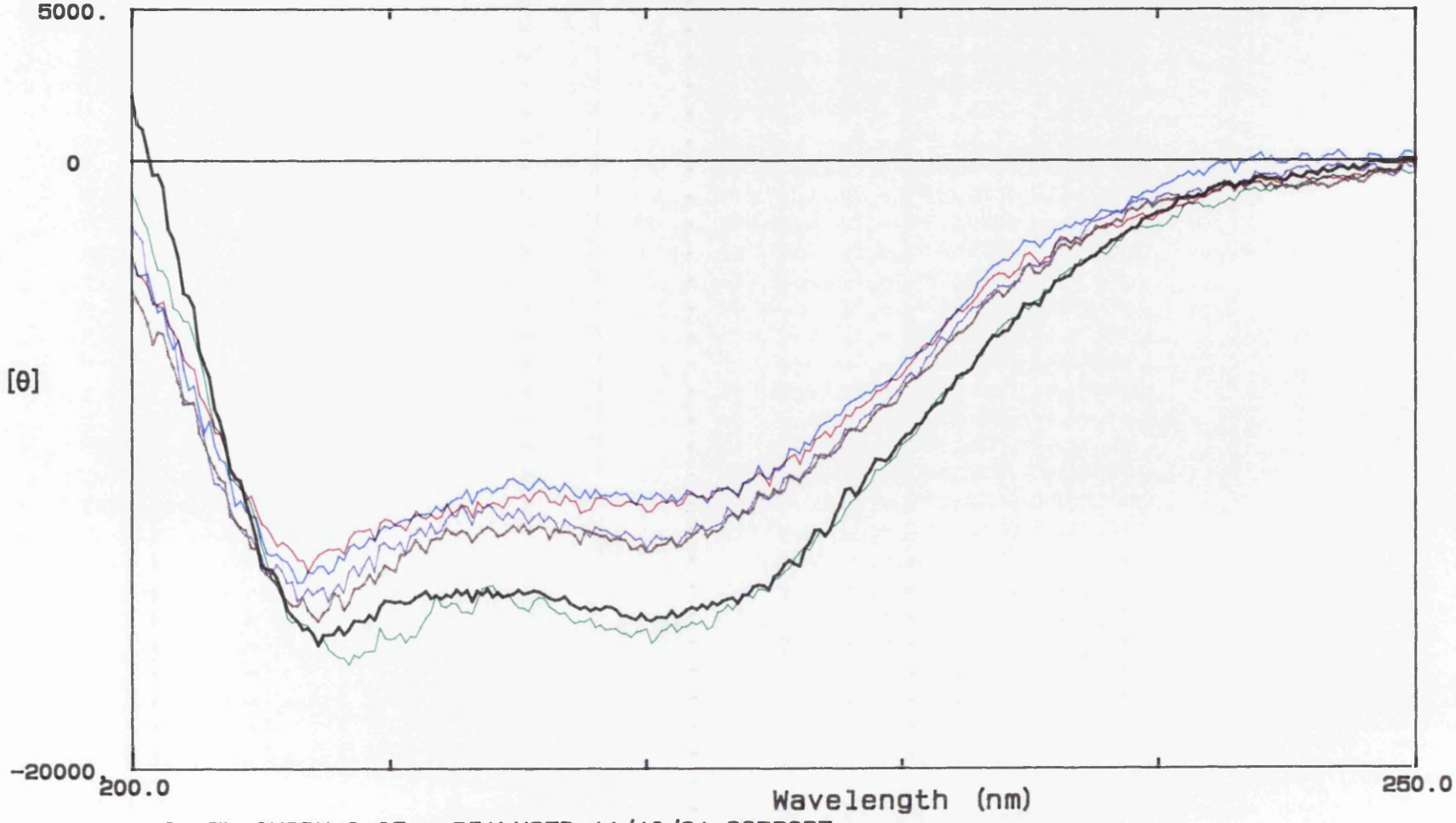
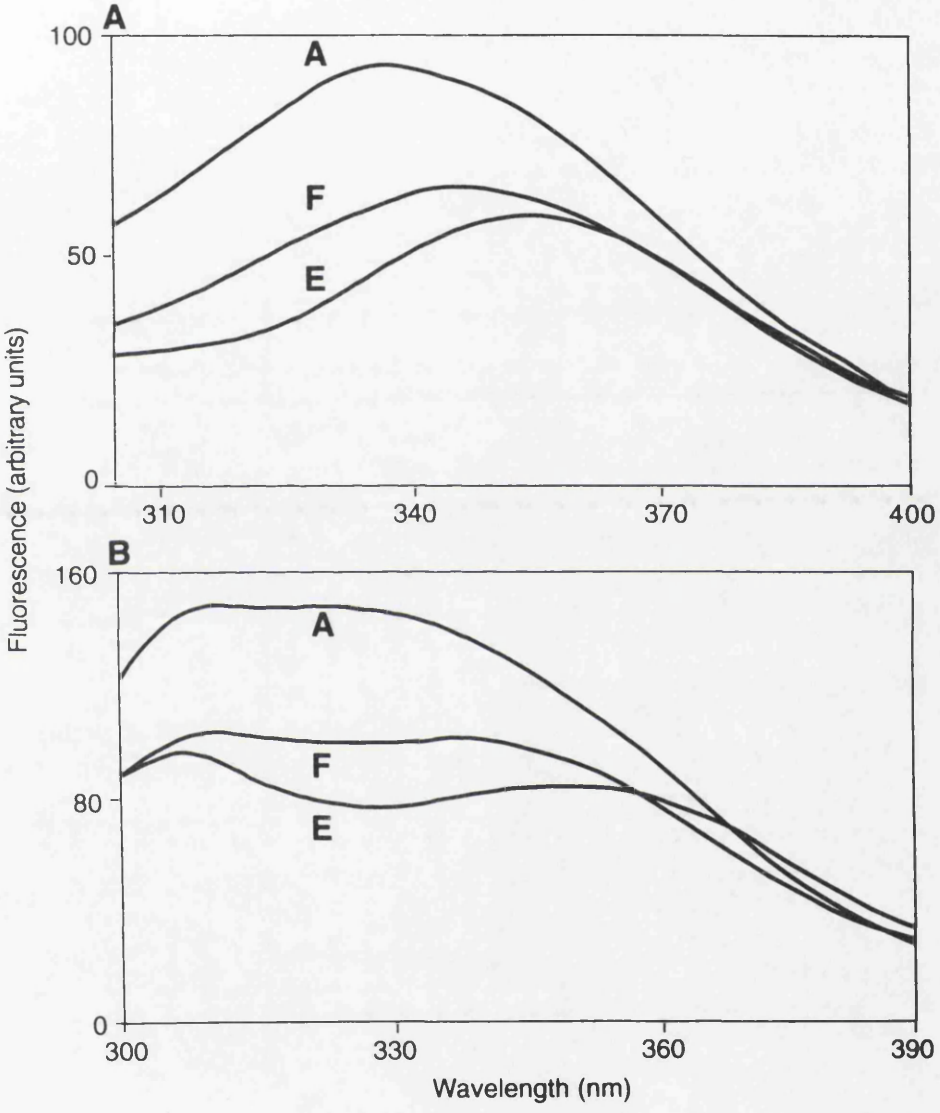

5000.
0
[θ]
-20000
200.0
250.0
Wavelength (nm)
GroEL 0MGDN 0.05cm DIALYSED 11/10/91 CORROPT
GroEL 0.5MGDN 0.05cm DIALYSED 11/10/91 CORROPT
GroEL 1.3MGDN 0.05cm DIALYSED 11/10/91 CORROPT
GroEL 1.5MGDN 0.05cm DIALYSED 11/10/91 CORROPT
GroEL 2MGDN 0.05cm DIALYSED 11/10/91 CORROPT
GroEL 6MGDN 0.05cm DIALYSED 11/10/91 CORROPT

Fig. 5.8 Changes in the fluorescence of GroEL after removal of GdnHCl by dialysis

Spectra were obtained in a buffer of 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 at 20°C. Unfolding of GroEL (50 μg / ml) in the presence of GdnHCl was carried out as described in section 2.8.2. GdnHCl was then removed by dialysis (section 2.8.3). Spectra A and E refer to GdnHCl concentrations of 0 and 6.0 M, respectively; spectrum F refers to the sample incubated in 6.0 M GdnHCl and then dialysed to remove the denaturing agent. **Fig. 5.8A** and **5.8B** refer to excitation at 290 nm and 280 nm, respectively.


A
100
A
F
E
50
0
310
340
370
400
B
160
A
F
E
80
0
300
330
360
390
Fluorescence (arbitrary units)
Wavelength (nm)

dialysed. It should be noted that on occasion the solutions became slightly turbid under these conditions (*i.e.* where the initial concentration of denaturant was greater than 1.0 M and then was removed by dialysis) suggesting that aggregate formation had occurred.

5.4 Characterisation of the 'intermediate' state of GroEL

GroEL was observed to regain secondary structure efficiently after unfolding, but the correct tertiary and quaternary structure and enzyme activity was not regained. In order to characterise this 'intermediate' state of GroEL further, and in particular to determine whether it possessed any properties of a 'molten-globule' folding intermediate (Ptitsyn *et al.* ,1990; Christensen and Pain, 1991), ANS binding and susceptibility to proteolysis were studied.

5.4.1 8-anilino-1-naphthalenesulphonate (ANS) binding

The 'molten globule' is an early folding state characterised by having most if not all of the secondary structure of the native protein in conjunction with a relatively compact, but unordered, flexible tertiary structure internalising a molten hydrophobic core (Ptitsyn *et al.* ,1990; Christensen and Pain, 1991). The fluorescence of the apolar probe ANS is dependent on the hydrophobicity of the environment and enhanced binding of the probe is suggestive of a 'molten globule' type structure.

The binding of ANS to native, unfolded and 'refolded' GroEL was measured by fluorescence as described in section 2.9.5. The fluorescence of ANS (20 μM) in the presence of 50 μg / ml GroEL was enhanced approximately 2.2 fold, compared to free ANS. This small enhancement declined further when GdnHCl was included over the range between 0 and 6.0 M (Fig. 5.9). The enhancement of only 1.3 fold (over free

ANS) at 6.0 M GdnHCl is consistent with the general observation that ANS does not bind to any great extent to either the native or fully unfolded states of proteins (Ptitsyn *et al.* ,1990; Christensen and Pain, 1991). The product obtained after dialysis of 6.0 M GdnHCl showed a small 2.1 fold enhancement fluorescence, compared with free ANS, suggesting that it does not posses a major characteristic of the 'molten-globule' state (*i.e.* enhanced binding of ANS). A 10 fold enhancement in ANS binding is typical for molten globules (Semisotnov *et al.*, 1987).

5.4.2 Susceptibility to thermolytic proteolysis

In order to study the compactness of the structure of the GroEL 'intermediate' state obtained by dialysis, limited proteolysis was carried out in the presence of thermolysin (section 2.9.6). Thermolysin was selected because (a), it is readily inactivated by treatment with EDTA, therefore preventing artefactual proteolysis during preparation for SDS-PAGE and (b), the preference for large hydrophobic side chains which occur only rarely at the surface of native proteins usually is an effective means for discriminating between native and non-native structures (Girg *et al.*, 1981).

Native GroEL was observed to be resistant to proteolysis with treatment with 2 - 100% (w / w) thermolysin as judged by SDS-PAGE (Fig. 5.10). By contrast, the product obtained after dialysis was very susceptible to proteolysis. After 10 minutes incubation with 2% (w / w) thermolysin only a small amount of material corresponding to an apparent M_r of 54 000 was observed (Fig. 5.10, lane 9). The intensity of this band corresponds to 28% of that observed for the intact polypeptide chain (M_r 57 000) in the absence of thermolysin (Fig. 5.10, lane 10) as judged by laser densitometric analysis. No distinct bands greater than 14 000 M_r were observed with larger amounts of protease. This data indicates that the dialysed sample has a more 'open' structure than the native protein.

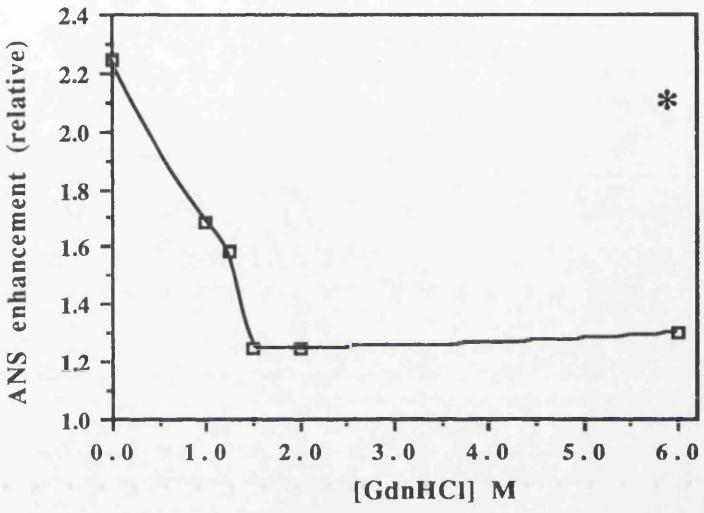


Fig. 5.9 ANS fluorescence of GdnHCl denatured-GroEL

Unfolding of GroEL (50 μg / ml) in the presence of various concentrations of GdnHCl was carried out as described in section 2.8.2, in a buffer of 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5 at 20°C. The fluorescence of ANS (20 μM) in the presence of GroEL was measured (section 2.9.5) and expressed in relative terms over the fluorescence of free ANS. (*) denotes the enhanced fluorescence of GroEL in 6.0 M GdnHCl which was then removed by dialysis.

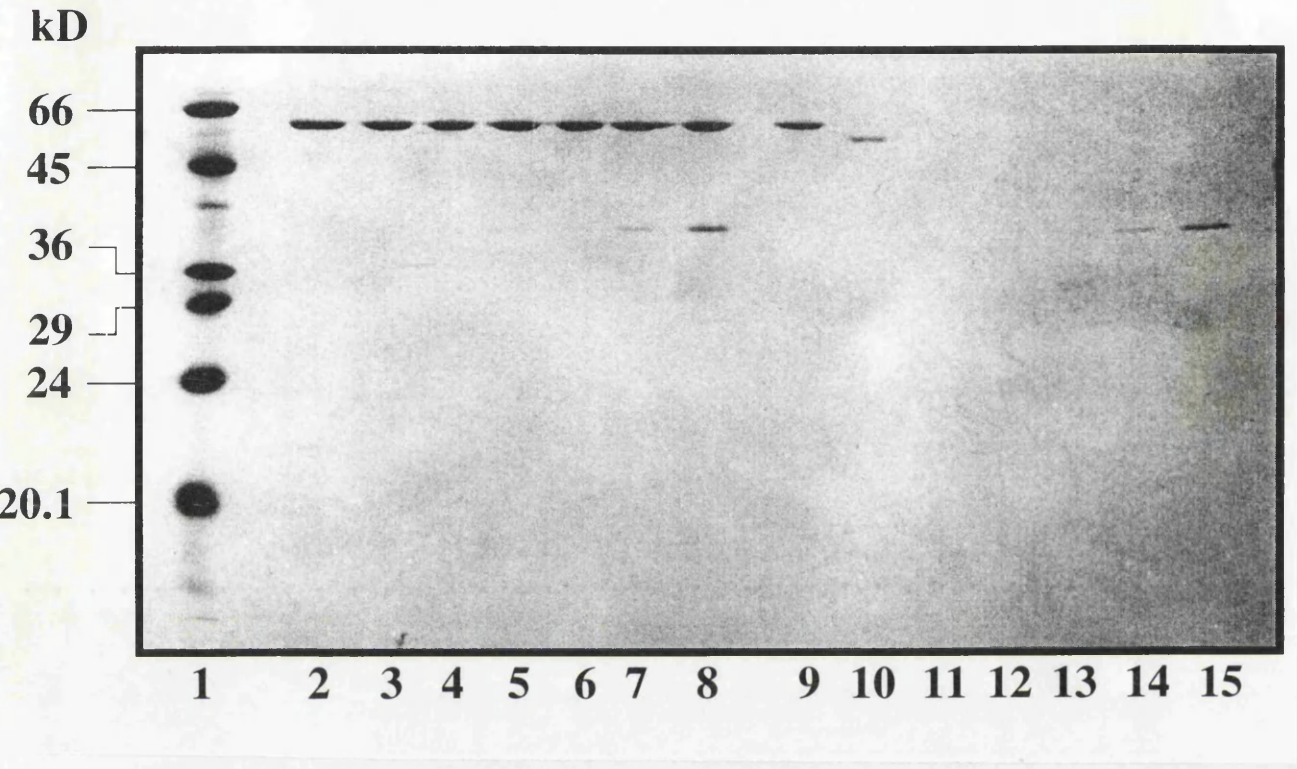


Fig. 5.10 Susceptibility of GroEL to proteolysis by thermolysin

GroEL (50 μg / ml) was incubated with thermolysin for 10 minutes at 20°C in 50 mM-potassium phosphate, 20mM-potassium acetate, 5 mM-magnesium acetate, 0.1 mM-dithiothreitol, pH 7.5. The reaction was stopped by the addition of 2 mM EDTA and samples analysed on 15% SDS-PAGE gels (section 2.3). Lanes 1, 2 - 8 and 9 - 15 contain M_r markers (section 2.10), native GroEL and GroEL which has been denatured and dialysed, respectively. The concentrations of thermolysin (μg / ml) added were: lanes 2 and 9, 0; lanes 3 and 10, 1; lanes 4 and 11, 2; lanes 5 and 12, 5; lanes 6 and 13, 10; lanes 7 and 14, 20; lanes 8 and 15, 50. The band at M_r 33 000 in lanes 7, 8, 14 and 15 correspond to thermolysin.

5.5 'Self-chaperoning' of GroEL in the presence of (Mg-ATP)

The proposal by Lissin *et al.* (1990) that urea dissociated-GroEL 14-mers can reassemble in the presence of (Mg-ATP) by virtue of a 'self-chaperoning' mechanism was examined with GdnHCl treated GroEL. GroEL (2.4 μM per protomer) was denatured in the presence of 6M GdnHCl as before (section 2.8.2) for 15 minutes at 20°C, then the denaturant removed by dialysis (section 2.8.3). The ATPase activity of this sample was less than 10% of the activity of a control in the absence of GdnHCl (as was previously observed in section 5.3.1). Incubation of the dialysed sample with 1 mM ATP for 20 minutes, showed a regain of ATPase activity of some 38% . Addition of native GroEL (0.6 μM per subunit) and 1 mM ATP to the dialysed sample for 20 minutes showed a slightly higher regain of ATPase activity of the order of 50% when compared to the control. These results are consistent with those of Lissin *et al.* (1990) and are suggestive of a regain of some tertiary and presumably quaternary structure of the dialysed protein. No structural studies were undertaken in conjunction with these experiments.

5.6 Discussion

The results reported here show that the denaturing agent guanidinium chloride (GdnHCl) causes major structural alterations in GroEL when the concentration is raised above about 1.0 M. The changes in secondary, tertiary and quaternary structure, as judged by circular dichroism, fluorescence and light scattering, respectively, run in parallel (Fig. 5.5). The steep loss of structure, without any discernible plateau, over a relatively narrow range of GdnHCl concentrations (1.0-1.5 M) points to a markedly cooperative unfolding of the tetradecamer. A similar cooperative behaviour has been reported for the unfolding of the GroEL homologue symbionin (a chaperonin present in the aphid endosymbiont), although this protein is more labile to disassembly than native GroEL in the presence of urea (Morioka and Ishikawa, 1993). This greater lability to

denaturation may reflect subtle differences in the oligomeric arrangements between the 2 proteins (symbionin also has phosphotransferase activity as well as ATPase activity). By way of contrast, the unfolding of the hexameric glutamate dehydrogenase occurs in two distinct stages via a trimeric intermediate as the GdnHCl concentration is increased (West and Price, 1988), indicating that complete unfolding in a cooperative fashion is not universal and presumably is dependent on the strength of intersubunit interactions. The ATPase activity of GroEL was lost at lower concentrations of GdnHCl (0.75 M) before any significant structural changes had occurred. It is likely that the more pronounced loss of ATPase activity is due to the greater sensitivity of the active site towards structural pertubations as has been observed for other proteins (Tsou, 1986; West and Price, 1988; West and Price, 1989; West and Price, 1990; Kelly and Price, 1991). An alternative explanation is that the denaturant itself is binding to important charged amino acid side-chains at the active site and one cannot rule out this possibility. It should be noted that because the unfolding process cannot be considered a reversible process (under these conditions), the changes in the various properties studied do not represent the attainment of a true equilibrium between the native and unfolded states of the protein. The protein should be regarded as reaching a metastable state, especially in the 'transition region' between 1.0-1.5 M-GdnHCl, which is observed to be stable over a period of 1 hour.

Attempts to refold the protein by removal of the denaturant by the dialysis technique proved to be largely unsuccessful. At concentrations at or below where the major structural alterations occurred (*i.e.* 1.0 M) there was essentially complete regain of secondary, tertiary and quaternary structure and in addition the ATPase activity was restored (Fig 5.6). When the initial concentration of GdnHCl was above the level where the major structural alterations occurred, GroEL was unable to regain a native-like conformation. In this respect the behaviour of GroEL is analogous to glutamate dehydrogenase (West and Price, 1988), aspartate aminotransferase (West and Price, 1989; West and Price, 1990), citrate synthase (West *et al.*, 1990) and fumarase (Kelly

and Price, 1991). Whilst the secondary structure of GroEL can be regained with high efficiency upon refolding from the completely denatured state (Fig. 5.7), the tertiary arrangement of the GroEL polypeptide upon dialysis is intermediate between that of the native and fully unfolded state (Fig. 5.8). The results obtained with the smallest amounts of thermolysin suggest that a folded structure involving most (54 kD out of 57 kD) of the polypeptide chain had been formed. However, this product still has a relatively open structure compared to native GroEL, since it is degraded by larger amounts of proteinase (Fig. 5.10). Not surprisingly, therefore, neither the ATPase activity nor the correct quaternary structure could be regained. The inability of GroEL to refold into its correct oligomeric conformation under the conditions used in this work is, at least in part, likely to reflect the size and complexity of the protein. A lack of proper coordination between the processes of polypeptide folding and domain pairing and thus oligomerisation may lead to the formation of a folding intermediate which is unable to attain its final native structure and may lead to aggregation of the protein (as was observed on occasion during these experiments). This type of behaviour has been observed for a number of multidomain and multisubunit proteins upon attempted refolding (see Price, 1992).

The inability of GroEL to reassemble correctly after denaturation poses interesting questions concerning the assembly of the protein *in vivo*. Lissin *et al.* (1990) have shown that the refolding of GroEL *in vitro* after denaturation in 4 M-urea requires the presence of Mg-ATP (and is enhanced by GroES) and have postulated the idea of 'self-chaperoning'. In this proposal GroEL monomers are capable of Mg-ATP dependent self-assembly and reformation of the native tetradecamer which in turn can stimulate the oligomerisation of the remainder of GroEL monomers in an ATPase dependent manner. This process occurs in a highly cooperative fashion. The results presented in this chapter also showed that a regain in the ATPase activity of GroEL could only be achieved by the presence of Mg-ATP (and was slightly enhanced by the addition of native GroEL), although no structural studies were undertaken with these

proteins. In a study of *mif* 4 mutants in yeast, Cheng *et al.* (1990) concluded that the assembly of the mitochondrial chaperonin Hsp60 *in vivo* is dependent on a pre-existant functional tetradecameric complex. Another GroEL homologue, symbionin, is required for its own assembly *in vitro* in an ATP-dependent process, analogous to GroEL (Morioka and Ishikawa, 1993). Thus these results, plus those described in this chapter, show that GroEL, and its homologues are unable to refold spontaneously . In the case of GroEL, and symbionin, refolding is dependent on the presence of Mg-ATP and is stimulated by native GroEL (and also GroES). Such requirements resemble the GroE mediated refolding of many polypeptides (section 1.5.6.4), thus GroEL not only chaperones the (re)folding of other proteins but also appears to chaperone its own assembly *in vitro*. As a point of interest it has been shown that a chemically synthesised version of the co-chaperonin GroES is able to assemble into its active oligomeric form in the absence of any co-factors (Mascagni *et al.*, 1991). Thus this protein differs from other oligomeric proteins such as rubisco in that it is able to spontaneously assemble in the absence of (other) chaperones.

The GroEL protein was observed to be less stable at lower temperatures, consistent with Lissin *et al.* (1990). GroEL monomers undergo cold denaturation at temperatures lower than 5°C and are unable to oligomerise into the native tetradecameric structure even in the presence of Mg-ATP (Lissin *et al.*, 1990; Lissin and Hemmingsen, 1993). The lack of stability at low temperatures would indicate that hydrophobic interactions play an important role in the stabilisation of the oligomer, since such interactions are less stable with decreasing temperatures. Recent structure-function studies of GroEL have indicated that the N-terminus of the polypeptide plays a crucial role in the assembly of the tetradecamer. Mutations in either Ala2 (Horovitz *et al.*, 1993a) or Lys3 (Horovitz *et al.*, 1993b) weaken intersubunit interactions and thus destabilise the oligomeric structure. A charge reversal of Lys3 to Glu completely prevents the oligomerisation of GroEL *in vivo* (Horovitz *et al.*, 1993b). The lysine residue at position 3 in the GroEL amino acid sequence is absolutely conserved

between the chaperonins and may therefore be critical for intersubunit stabilisation. The stability of the native GroEL tetradecamer is therefore likely to involve multiple intersubunit interactions involving both charged and hydrophobic interactions. Interestingly, although the N-terminal of GroEL is crucial for its correct assembly the conserved glycine and methionine-rich motif found at the C-terminus appears to be dispensable. Removal of these 16 C-terminal residues do not affect the ability of GroEL to function either *in vitro* or *in vivo* (Langer *et al.*, 1992b; McLennan *et al.*, 1993). In order to completely understand the role of GroEL, the structural basis of its own assembly requires to be fully elucidated. It remains to be directly demonstrated that GroEL is able to 'self-chaperone' *in vivo* and / or whether additional, as yet unidentified, chaperone-like proteins exist which may (also) assist in the assembly of the GroEL tetradecamer.

To summarise the results of this chapter, GroEL monomers, in the absence of ATP, are able to spontaneously regain secondary structure but not higher order structure or ATPase activity. The product obtained upon removal of denaturant is more sensitive to proteolysis by thermolysin, unlike the native protein, but does not show enhanced binding of ANS, a characteristic property of the 'molten-globule' state of proteins. Regain of native-like properties, as judged by regain of ATPase activity, requires the presence of ATP and is enhanced by native GroEL. It remains a task for future work to characterise this 'folded intermediate' state of GroEL in more detail and to further explore the role played by GroES and ATP in the later steps involved in the acquisition of the correct tertiary and quaternary structures of the GroEL protein.

Chapter 6 General discussion and future prospects

6.1 Introduction

The aims of the work presented in this thesis was to study structure-function relationships of the GroEL protein with regards to (1) the location of the functionally important ATP binding site and (2) the nature of the GroEL (self-)assembly process. The preceding chapters have described in detail the progress made in fulfilling these original objectives and are summarised below.

6.2 The ATP binding site of GroEL

The binding and hydrolysis of ATP by GroEL is the critical step for the release of 'polypeptide substrates' from the chaperone protein, yet when this work was initiated nothing at all was known about the location of or the important amino acid residues present at the ATP binding site. Using an affinity labelling approach, residues towards the C-terminus of GroEL were identified as being present at or near the ATP binding site of GroEL (Chapter 4). In particular two residues Cys457 and / or Cys518 were proposed to be the site of reaction for the reactive aldehyde group(s) of the ATP analogue 2' 3' dialdehyde ATP (oATP) (published in Thomson *et al.*, 1993). This proposal is consistent with other affinity labelling approaches undertaken and since published (Martin *et al.*, 1993a; Bochkareva *et al.*, 1994 :- discussed in detail in section 4.9). Although individual reactive groups have now been identified which are presumably present in the location of the ATP binding site, it is unknown whether they are absolutely essential for the ATPase activity of GroEL. The 3 cysteine residues present in GroEL are not conserved, possibly indicating a non-essential role. Essential, and conserved, amino acid side chains such as those of lysine and arginine, often found in the phosphate binding regions of nucleotide binding sites have not been identified to date. In the studies undertaken in this work, lysine specific reagents such as pyridoxal 5'-phosphate and methylacetimidate were ineffective in inhibiting the GroEL ATPase activity, under the conditions employed. Identification of these catalytically important

residues are likely to come, if not from chemical modification then, from X-ray crystallography studies. To date GroEL crystals which diffract to a low resolution have been obtained (Spangfort *et al.*, 1993; Svensson *et al.*, 1994). These crystals diffract only to a maximum of 7 Å (Spangfort *et al.*, 1993) and it will take a higher resolution diffraction before the three dimensional structure of GroEL and thus the molecular mechanism of its function is solved.

It will also be of great interest to know whether the limited sequence conservation between the chaperonins and their weakly related homologues in the eukaryotic cytosol (the heterologous TRiC class) reflects the maintenance of ATPase activity (Lewis *et al.*, 1992) or whether the ATP binding pockets of these two classes of protein are similar but non identical. The latter possibility would be consistent with the putative ATP triphosphate binding regions of GroEL identified by Martin *et al.* (1993b) which shows homology to other members of the 'classical' chaperonin class of proteins but not to members of the TRiC class. It is known that at least two classes of purine nucleotide triphosphate binding folds exist. Those proteins with the Walker *et al.* (1982) type A motif (e.g. adenylate kinase, elongation factor Tu, myosin) have a common nucleotide binding fold (Saraste *et al.*, 1990; Milner-White *et al.*, 1991) which is distinct from other ATP binding proteins (e.g. Hsp70s, actin, hexokinase) (Bork *et al.*, 1992). We must wait for crystallography studies to determine the nature of the ATP binding fold(s) within the chaperonin and TRiC classes.

6.3 GroEL assembly

The question of how GroEL itself assembles into it characteristic tetradecameric structure is also one which has not been fully resolved. The results reported in chapter 5 (published in Price *et al.*, 1993) indicate that the GroEL protein is unable to reassemble into its native structure following denaturation in guanidinium chloride. Unfolding of the protein appears to be a highly cooperative process, although the amino

acids present at the ATPase active site appear to be more sensitive to small structural alterations. Although the polypeptide can spontaneously regain secondary structure upon removal of denaturant, regain of functional chaperonin properties requires the presence of Mg-ATP and is enhanced by native GroEL. These results are consistent with the proposal that GroEL and its homologues can 'self-chaperone' *in vitro* (Lissin *et al.*, 1990; Morioka and Ishikawa, 1993). It still remains for the folding pathway of the GroEL protein to be characterised in more detail and to determine whether GroEL can self-assemble *in vivo*. The *in vivo* assembly of the GroEL homologue in yeast mitochondria, Hsp60, is dependent on the presence of a pre-existing functional Hsp60 complex (Cheng *et al.*, 1990). This protein is nuclear encoded and then transported into the mitochondrial matrix. It is therefore likely that chaperones of the Hsp70 class are involved at some stage in the translocation of the polypeptide into the matrix and its subsequent assembly into native structure, although, as yet, this has not been demonstrated. A sequential folding pathway involving the *E. coli* Hsp70 homologue DnaK and GroEL has also been proposed for the assembly of some prokaryotic cytosolic proteins (Langer *et al.*, 1992a), therefore it is feasible that DnaK and its co-chaperones may also play an important role in the assembly of the GroEL chaperonin *in vivo*. Thus whilst GroEL (and its homologues) may indeed undergo a process of 'self-chaperoning' *in vivo*, one cannot rule out the involvement of additional (possibly as yet unidentified) factors in the attainment of native structure. It should also be noted that of the experimental criteria used to indicate that both symbionin and GroEL undergo 'self-chaperoning' into their native oligomeric form *in vitro* (*i.e.* as judged by molecular mass, regain of ATPase activity and electron microscopic appearance) no reconstitution studies using reassembled chaperonins and other model protein systems have been undertaken. It remains to be demonstrated, for example, whether reassembled GroEL can also assist in the reassembly of denatured rubisco in a manner analogous to native GroEL. This point needs to be clarified in order to confirm that the correct native-like three dimensional structure of GroEL has been obtained.

6.4 Unanswered questions in the field of chaperone proteins

The implication that chaperone proteins play important roles in general *de novo* protein folding is one that has come about only relatively recently. Although research in this field is progressing rapidly the stuctural basis and mechanism of chaperone function is not fully understood. Some of the oustanding questions in this field will be outlined below.

The structural basis of chaperonin recognition of polypeptide substrates in not yet fully understood. The available evidence would suggest that GroEL recognises hydrophobic regions present in early folding intermediates of several (unrelated) polypeptides which also contain some elements of secondary structure (discussed in section 1.5.6.6). Whether or not these early folding intermediates contain some sort of target or consensus sequence which may be recognised by GroEL has yet to be firmly elucidated , as has the nature of the GroEL polypeptide binding 'pocket' itself. Martin *et al.* (1993a) proposed a reaction mechanism of the GroEL-GroES cycle whereby the polypeptide substrate bound within the central cavity, presumably at multiple sites within the channel to allow a cooperative release, and was released in a folding competent state by the concerted actions of ATP and GroES. These authors proposed that the unfolded polypeptide initially bound to the available free GroEL ring opposite to the ring with bound GroES. Electron microscopic analysis of the *Thermus thermophilus* holo-chaperonin (*i.e.* a homologue of the GroEL-GroES complex) has recently shown that the unfolded substrate polypeptide 3-isopropylmalate dehydrogenase does indeed bind to the opposite end of chaperonin-60 to where chaperonin-10 is bound (Isii *et al.*, 1994). This is likely, therefore, to be a common mechanism between all polypeptides which undergo GroE mediated folding *in vivo*. This raises some more interesting questions about the role of the GroE proteins during chaperonin-mediated folding, which at present are unresolved. For example, during multiple cycles of polypeptide binding and release do (some) folding polypeptides

shuttle between the two rings of GroEL or do they interact only with the same ring they initially bind to and, therefore, does GroES (also) shuttle between both ends of the GroEL cylinder to coordinate the release of the polypeptide substrate ?. Surprisingly GroES has also been shown to be able to bind to but not hydrolyse ATP (Martin *et al.*, 1993b). An additional role of GroES has therefore been proposed to donate ATP to GroEL and thus increase the cooperativity of ATP binding which presumably would increase the efficiency of GroE-mediated folding. This interesting proposal remains to be confirmed although it should be relatively simple to demonstrate any transfer of adenine nucleotides between the chaperonins.

One of the other outstanding questions is does a sequential chaperone folding pathway, involving Hsp70s and Hsp60s, also exist for the folding of polypeptides within the cytosol, as has been demonstrated using purified *E. coli* chaperone proteins *in vitro* (Langer et al., 1992a). This pathway (which exists in mitochondria, and is described in section 1.6) would potentially minimise the chances of aggregation or misfolding of many, and especially larger, polypeptides but it remains to be confirmed *in vivo* and demonstrated in cell-free translation systems *in vitro*. It seems likely that such a folding pathway does exist in the cytosol, at least for some polypeptides, in view of the differential binding capabilities of the two classes of chaperone proteins. Whether such a folding pathway exists in the eukaryotic cytosol remains unclear to date, in view of the relative novelty of the eukaryotic cytosolic chaperonin-like proteins (the TRiC class).

6.5 Evolution of the chaperonin family

Studies on the *E. coli* GroE proteins has given a remarkable insight into how many, but not all, polypeptides require their assistance to attain native structure both *in vitro* and *in vivo*. The remarkable structural and functional conservation between GroEL and GroES and their analogues in mitochondria and chloroplasts has given an

indication of the importance of these proteins within the cell and indeed these proteins are essential for normal cellular function. These proteins probably evolved from a common eubacterial origin. No homologues have, however, been identified in the eukaryotic cytosol to date. Their closest relations appear to be the weakly related members of the TRiC sub-family of chaperonins which have been found to be present in the cytosol of all eukaryotic cells examined to date (see Frydman *et al.*, 1992; Gao *et al.*, 1992; Lewis *et al.*, 1992; Yaffe *et al.*, 1992). Like the 'classical' chaperonins these proteins have a double-ring structure which has ATPase activity, unlike them, however, the stacked ring structure exists as hetero-oligomeric complex arranged with an eight or nine fold symmetry. The heteromeric nature of the TRiC complex consists of up to ten different subunits several of which, but not all, are related to the ubiquitous eukaryotic protein TCP1 (Frydman *et al.*, 1992; Kubota *et al.*, 1994). The diversity of subunits present in the complex may give it a greater spectrum of potential polypeptide substrates (compared to 'classical' chaperonins) and reflect the evolution of more complex functions required to deal with the folding and assembly more highly evolved polypeptides present in the eukaryotic cytosol. The TRiC complex has already been shown to facilitate the (ATP assisted) folding of tubulin, actin and firefly luciferase *in vitro* (Frydman *et al.*, 1992; Gao *et al.*, 1992; Yaffe *et al.*, 1992). It is likely that other polypeptides functionally interact with the TRiC complex and that it will have a general role in polypeptide folding within the eukaryotic cytosol, possibly equivalent to the 'classical' chaperonins. The common structural and functional similarities of the chaperonins and the TRiC sub-family of chaperonins suggests that they either arose from a common ancestor or evolved convergently to carry out chaperoning functions in their respective compartments. Studies on the TRiC complex are still in its infancy and a lot is unknown about its function. No chaperonin-10 like co-factor (if it exists) has been identified for TRiC and it still remains to be directly demonstrated that the complex assists polypeptide folding *in vivo*. The bacterial GroE system may provide important clues about its mechanism of action.

6.6 Future prospects

It seems likely that the list of potential chaperone proteins will grow longer as further interest and research in this field is generated. Indeed a putative chaperonin 60 protein has recently been identified, after heat shock, in the nucleus of a fish cell line (Sanders *et al.*, 1994) whilst bacteriophage T4 encodes a protein which, despite absence of sequence similarity, can substitute for GroES in the GroEL dependent refolding of rubisco *in vitro* and *in vivo* and assist bacterial growth at non-permissive temperatures (van der Vies *et al.*, 1994). These and other findings, including the recent proposals that both protein disulphide isomerase and peptidyl prolyl *cis-trans* isomerase have chaperone activities (discussed in sections 1.4.1 and 1.4.2), may indicate the need for the definition and / or the categorisation of chaperone proteins to be altered. The danger is that some proteins will be proposed as chaperones by virtue of simply providing an interactive surface which may have the ability to act as a 'biochemical sponge' and so prevent aggregation of a given polypeptide in the *in vitro* situation but will not have this function *in vivo*. These situations must be clarified in more detail.

The GroE chaperone system is now relatively well characterised and it remains for eukaryotic chaperonins (in particular the TRiC class) to be further explored. Some of the outstanding questions regarding chaperonin mediated protein folding have been outlined above, and will no doubt be source of future research in order to gain further insights into the mechanism of how these remarkable components assist in the acquisition of correct protein structure within the cell. In a practical sense it is likely that co-expression of these proteins will provide to be of great use to maximise expression of recombinant proteins both in bacteria and in transgenic plants. For medical interest it may be of importance to identify dysfunctional chaperone proteins which may (possibly) affect protein folding and thus have implications which may lead to disease.

Providing the answers to such questions will, no doubt, prove to be of great significance

References

Anfinsen, C.B. (1973) *Science* **181**, 223-230

Badcoe, I.G., Smith, C.J., Wood, S., Halsall, D.J., Holbrook, J., Lund, P. and Clarke, A.R. (1991) *Biochemistry* **30**, 9195-9200

Baker, D., Sohl, J.L. and Agard, D.A. (1992) *Nature* **356**, 263-265

Baneyx, F. and Gatenby, A.A. (1992) *J. Biol. Chem.* **267**, 11637-11644

Bardwell, J.C.A. and Craig, E.A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 848-852

Bardwell, J.C. and Craig, E.A. (1988) *J. Bacteriol.* **170**, 2977-2983

Barraclough, R. and Ellis, R.J. (1980) *Biochim. Biophys. Acta.* **608**, 19-31

Bächinger, H.P. (1987) *J. Biol. Chem.* **262**, 17144-17148

Bächinger, H.P. and Compton, L.A. (1991) *J. Cell. Biochem. suppl.* **15G**, 188

Bergman, L.W. and Kuehl, W.M. (1979) *J. Biol. Chem.* **254**, 8869-8876

Bernikov, L.R., Dzhandzhugazyan, K.N., Lutsenko, S.V. and Modyanov, N.M. (1990) *Eur. J. Biochem.* **194**, 413-421

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

Blond-Elguindi, S., Cwirla, S.E., Dower, W.J., Lipshutz, R.J., Sprang, S.R., Sambrook, J. and Gething, M.-J.H. (1993a) *Cell* **75**, 717-728

Blond-Elguindi, S., Fourie, A.M., Sambrook, J. and Gething, M.-J.H. (1993b) *J. Biol. Chem.* **268**, 12730-12735

Bochkareva, E.S., Lissin, N.M., Flynn, G.C., Rothman, J.E. and Girshovich, A.S. (1992) *J. Biol. Chem.* **267**, 6796-6800

Bochkareva, E.S., Horovitz, A. and Girshovich, A.S. (1994) *J. Biol. Chem.* **269**, 44-46

Bork, P., Sander, C. and Valencia, A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7290-7294

Borkovich, K., Farrelly, F., Finkelstein, D., Taulien, J. and Lindquist, S. (1989) *Mol. Cell. Biol.* **9**, 3919-3930

Bradford, M.M. (1976) *Anal. Biochem.*. **72**, 248-254

Braig, K., Simon, M., Furuya, F., Hainfeld, J.F. and Horwich, A.L. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3978-3982

Brandsch, R., Bichler, V., Schmidt, M. and Buchner, J. (1992) *J. Biol. Chem.* **267**, 20844-20849

Bresnick, E.H., Dalman, F.C., Sanchez, E.R. and Pratt, W.B. (1989) *J. Biol. Chem.* **264**, 4992-4997

Brown, J.H., Jardetzky, T.S., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L. and Wiley, D.C. (1993) *Nature* **364**, 33-39

Buchner, J., Schmidt, M., Fuchs, M., Jaenicke, R., Rudolph, R., Schmid, F.X. and Kiefhaber, T. (1991) *Biochemistry* **30**, 1586-1591

Bukau, B. and Walker, G.C. (1989) *J. Bacteriol.* **171**, 2337-2346

Bulleid, N.J. and Freedman, R.B. (1988) *Nature* **335**, 649-651

Carrillo, N., Ceccarelli, E.A., Krapp, A.R., Boggio, S., Ferreyra, R.G. and Viale, A.M. (1992) *J. Biol. Chem.* **267**, 15537-15541

Chandrasekhar, G.N., Tilly, K., Woolford, C., Hendrix, R. and Georgopolous, C. (1986) *J. Biol. Chem.* **261**, 12414-12419

Chappell, T.G., Konforti, B.B., Schmid, S.L. and Rothman, J.E. (1987) *J. Biol. Chem.* **262**, 746-751

Cheng, M.Y., Hartl, F.-U, Martin, J., Pollock, R.A., Kalousek, F., Neupert, W., Hallberg, E.M., Hallberg, R.L. and Horwich, A.L. (1989) *Nature* **337**, 620-625

Cheng, M.Y., Hartl, F.-U. and Horwich, A.L. (1990) *Nature* **348**, 455-458

Christensen, H. and Pain, R. H. (1991) *Eur. Biophys. J.* **19**, 221-229

Colman, R. F., Pal, P. K. and Wyatt, J. L. (1977) *Methods Enzymol.* **46**, 240-249

Colman, R.F. (1983) *Ann. Rev. Biochem.* **52**, 67-91

Craig, E.A. and Gross, C.A. (1991) *Trends Biochem. Sci.* **16**, 135-140

deJong, W.W, Leunissen, J.A., Leenen, P.J., Zweers, A. and Versteeg, M. (1988) *J. Biol. Chem.* **263**, 5141-5149

Dingwall, C. and Laskey, R.A. (1990) *Sem. Cell Biol.* **1**, 11-17

Easterbrook-Smith, S.B., Wallace, J.C. and Keech, D.B. (1976) *Eur. J. Biochem.* **62**, 125-130

Edgerton, M.D., Santos, M.O. and Jones, A.M. (1993) *Plant Mol. Biol.* **21** 1191-1194

Edman, J.C., Ellis, L., Blacher, R.W., Roth, R.A. and Rutter, W.J. (1985) *Nature* **317**, 267-270

Ellis, R.J. (1990) *Sem. Cell Biol.* **1**, 1-9

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

Ellman, G.L. (1959) *Arch. Biochem. Biophys.* **82**, 70-77

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

Felsenfeld, G. (1978) *Nature* **271**, 115-122

Finley, D., Bartel, B. and Varshavsky, A. (1989), *Nature* **338**, 394-401

Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T. and Schmid, F.X. (1989) *Nature* **337**, 476-478

Fisher, M.T. (1992) *Biochemistry* **31**, 3955-3963

Flaherty, K.M., DeLuca-Flaherty, C. and McKay, D.B. (1990) *Nature* **346**, 623-628

Flaherty, K.M., McKay, D.B., Kabsch, W. and Holmes, K.C. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5041-5045

Flynn, G.C., Chappell, T.G. and Rothman, J.E. (1989) *Science* **245**, 385-390

Flynn, G.C., Pohl, J., Flocco, M.T. and Rothman, J.E. (1991) *Nature* **353**, 726-730

Flynn, G.C., Beckers, C.J.M., Baase, W.A. and Dahlquist, F.W. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10826-10830

Freedman, R.B. (1989) *Cell* **57**, 1069-1072

Freskgard, P.O., Bergenhem, N., Jonsson, B.H., Svensson, M. and Carlsson, U. (1992) *Science* **258**, 466-468

Friedman, D.E., Olsen, E.R., Georgopolous, C., Tilly, K., Herskowitz, I. and Banuett, F. (1984) *Microbiol. Rev.* **48**, 299-325

Fry, D.C., Kuby, S.A. and Mildvan, A.S. (1985) *Biochemistry* **24**, 4680-4684

Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J.S., Tempest, P. and Hartl, F.-U. (1992) *EMBO J.* **11**, 4767-4778

Gao, Y., Thomas, O., Chow, R.L., Lee, G.-H. and Cowan, N.J. (1992) *Cell* **69**, 1043-1050

Georgopoulos, C. P., Hendrix, R.W., Kaiser, A.D. and Wood, W.B. (1972) *Nature* **239**, 38-42

Georgopoulos, C. P., Hendrix, R.W., Casjens, S.R. and Kaiser, A.D. (1973) *J. Mol. Biol.* **76**, 45-60

Georgopoulos, C. (1992) *Trends Biochem. Sci.* **17**, 295-299

Gething, M.-J. and Sambrook, J. (1992) *Nature* **355**, 33-45

Girg, R., Rudolph, R. and Jaenicke, R. (1981) *Eur. J. Biochem.* **119**, 301-305

Goloubinoff, P., Christeller, J.T., Gatenby, A.A. and Lorimer, G.H. (1989) *Nature* **342**, 884-889

Gragerov, A., Nudler, E., Komissarova, N., Gaitanaris, G.A., Gottesman, M.E. and Nikiforov, V. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10341-10344

Gragerov, A., Zeng, L., Zhao, X., Burkholder, W. and Gottesman, M.E. (1994) *J. Mol. Biol.* **235**, 848-854

Gray, T.E. and Fersht, A.R. (1991) *FEBS Lett.* **292**, 254-258

Gray, T.E. and Fersht, A.R. (1993) *J. Mol. Biol.* **232**, 1197-1207

Grimm, R., Donaldson, G.K., van der Vies, S.M., Schäfer, E. and Gatenby, A.A. (1993) *J. Biol. Chem.* **268**, 5220-5226

Hallberg, R.L. (1990) *Sem. Cell Biol.* **1**, 37-45

Harding, M.W., Galat, A., Uehling, D.E. and Schreiber S.L. (1989) *Nature* **341**, 758-760

Hartman, D.J., Hoogenraad, N.J., Condron, R. and Høj, P.B.(1992) *Proc. Natl. Acad. Sci. USA* **89**, 3394-3398

Hartman, D.J., Surin, B.P., Dixon, N.E., Hoogenraad, N.J. and Høj, P.B. (1993a) *Proc. Natl. Acad. Sci. USA* **90**, 2276-2280

Hartman, D.J., Hoogenraad, N.J., Condron, R. and Høj, P.B.(1993b) *Biochim. Biophys. Acta* **1164**, 219-222

Hayer-Hartl, M.K. and Hartl, F.U. (1993) *FEBS Lett.* **320**, 83-84

Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K., Dennis, D.K., Georgopoulos, C., Hendrix, R.W. and Ellis, R.J. (1988) *Nature* **333**, 330-334

Hendrick, J.P. and Hartl, F.-U. (1993) *Annu. Rev. Biochem.* **62**, 349-384

Hendrick, J.P., Langer, T., Davis, T.A., Hartl, F.-U. and Wiedmann, M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10216-10220

Hendrix, R.W. (1979) *J. Mol. Biol.* **129**, 375-392

Hohn, T., Hohn, B., Engel, A. and Wurtz, M. (1979) *J. Mol. Biol.* **129**, 359-373

Holmgren, A. and Bränden, C.- I. (1989) *Nature* **342**, 248-251

Horovitz, A., Bochkareva, E.S., Kovalenko, O. and Girshovich, A.S. (1993a) *J. Mol. Biol.* **231**, 58-64

Horovitz, A., Bochkareva, E.S. and Girshovich, A.S. (1993b) *J. Biol. Chem.* **268**, 9957-9959

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

Horwitz, J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10449-10453

Höll-Neugebauer, B., Rudolph, R., Schmidt, M. and Buchner, J. (1991) *Biochemistry* **30**, 11609-11614

Ishii, N., Taguchi, H., Sumi, M. and Yoshida, M. (1992) *FEBS Lett.* **299**, 169-174

Ishii, N., Taguchi, H., Sasabe, H. and Yoshida, M. (1994) *J. Mol. Biol.* **236**, 691-696

Jackson, G.S., Stainforth, R.A., Halsall, D.J., Atkinson, T., Holbrook, J.J., Clarke, A.R. and Burston, S.G. (1993) *Biochemistry* **32**, 2554-2563

Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol.* **49**, 117-237

Jakob, U., Gaestel, M., Engel, K. and Buchner, J. (1993) *J. Biol. Chem.* **268**, 1517-1520

Jenkins, A.J., March, J.B., Oliver, I.R. and Masters, M. (1986) *Mol. Gen. Genet.* **202**, 446-454

John, D.C.A., Grant, M.E. and Bulleid, N.J. (1993) *EMBO J.* **12**, 1587-1595

Johnson, C.M. and Price, N.C. (1988) *Biochem. J.* **252**, 111-117

Kang, P.-J., Ostermann, J., Shilling, J., Neupert, W., Craig, E.A. and Pfanner, N. (1990) *Nature* **348**, 137-143

Kelly, S.M. and Price, N.C. (1991) *Biochem. J.* **275**, 745-749

King, M.M. and Carlson, G.M. (1981) *Biochemistry* **20**, 4382-4387

King, M.M. and Colman, R.F. (1983) *Biochemistry* **22**, 1656-1665

Kitz, R. and Wilson, I.B. (1962) *J. Biol. Chem.*. **237**, 3245-3249

Koivu, J. and Myllylä, R. (1987) *J. Biol. Chem.* **262**, 6159-6164

Kubo, T., Mizobata, T. and Kawata, Y. (1993) *J. Biol. Chem.* **268**, 19346-19351

Kubota, H., Hynes, G., Carne, A., Ashworth, A. and Willison, K. (1994) *Curr. Biol.* **4**, 89-99

Kumar, G., Kalra, V.K. and Brodie, A.F. (1979) *J. Biol. Chem.* **254**, 1964-1971

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

Laminet, A.A., Ziegelhoffer, T., Georgopoulos, C. and Plückthun, A. (1990) *EMBO J.* **9**, 2315-2319

Landry, S.J. and Gierasch, L.M. (1991) *Biochemistry* **30**, 7359-7362

Landry, S., Jordan, R., McMacken, R. and Gierasch, L. (1992) *Nature* **355**, 455-457

Landry, S.J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C. and Gierasch, L.M. (1993) *Nature* **364**, 255-258

Lang, K., Schmid, F.X. and Fischer, G. (1987) *Nature* **329**, 268-270

Lang, K. and Schmid, F.X. (1988) *Nature* **331**, 453-455

Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K.and Hartl, F.-U. (1992a) *Nature* **356**, 683-689

Langer, T., Pfeiffer, G., Martin, J., Baumeister, W. and Hartl, F.U. (1992b) *EMBO J.* **11**, 4757-4765

Laskey, R.A., Honda, B.M., Mills, A.D.and Finch, J.T. (1978) *Nature* **275**, 416-420

Lee, S.C. and Olins, P.O. (1992) *J. Biol. Chem.* **267**, 2849-2852

Lecker, S., Lill, R., Ziegelhofer, T., Georgopoulos, C., Bassford, P.J., Kunomoto, C.A. and Wickner, W. (1989) *EMBO J.* **8**, 2703-2709

Levinthal, C. (1968) *J. Chim. Phys.* **65**, 44-45

Levy, H.M., Leber, P.D. and Ryan, E.M. (1963) *J. Biol. Chem.*. **238**, 3654- 3659

Lewis, V.A., Hynes, G.M., Zheng, D., Saibil, H. and Willison, K. (1992) *Nature* **358**, 249-252

Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C. and Zylicz, M. (1991a) *Proc. Natl. Acad. Sci. USA* **88**, 2874-2878

Liberek, K., Skowyra, D., Zylicz, M., Johnson, C. and Georgopoulos, C. (1991b) *J. Biol. Chem.* **266**, 14491-14496

Lissin, N.M., Venyaminov, S. Y. and Girshovich, A. S. (1990) *Nature* **348**, 339-342

Lissin, N.M. and Hemmingsen, S.M. (1993) *FEBS Lett.* **324**, 41-44

Lubben, T.H., Gatenby, A.A., Donaldson, G.K., Lorimer, G.H. and Viitanen, P.V. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7683-7687

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L. and Hartl, F.-U. (1991) *Nature* **352**, 36-42

Martin, J., Mayhew, M., Langer, T. and Hartl, F.-U. (1993a) *Nature* **366**, 228-233

Martin, J., Geromanos, S., Tempest, P. and Hartl, F.-U. (1993b) *Nature* **366**, 279-282

Mascagni, P., Tonolo, M., Ball, H., Lim, H., Ellis, R.J. and Coates, A. (1991) *FEBS Lett.* **286**, 201-203

Matsudaria, P.(1987) *J. Biol. Chem.*. **262**, 10035-10038

McLennan, N.F., Girshovich, A.S., Lissin, N.M., Charters, Y. and Masters, M. (1993) *Mol. Microbiol.* **7**, 49-58

McMullen, T.W. and Hallberg, R.L. (1987) *Mol. Cell. Biol.* **7**, 4414-4423

Mendoza, J.A., Rogers, E., Lorimer, G.H. and Horowitz, P.M. (1991) *J. Biol. Chem.* **266**, 13044-13049

Mendoza, J.A. and Horowitz, P.M. (1992) *J. Protein Chem.* **11**, 589-594

Mendoza, J.A., Butler, M.C. and Horowitz, P.M. (1992) *J. Biol. Chem.* **267**, 24648-24654

Merck, K.B., Groenen, P.T.T.A., Voorter, C.E.M., deHaard-Hoekman, W.A., Horwitz, J., Bloemendal, H. and deJong, W.W. (1993) *J. Biol. Chem.* **268**, 1046-1052

Miller, A.D., Maghlaoui, K., Albanese, G., Kleinjan, D.A. and Smith, C. (1993) *Biochem. J.* **291**, 139-144

Mills, A.D., Laskey, R.A., Black, P. and De Robertis , E.M. (1980) *J. Mol. Biol.* **139**, 561-568

Milner-White, E.J., Coggins, J.R. and Anton, I.A. (1991) *J. Mol. Biol.* **221**, 751-754

Mizobata, T., Akiyama, Y., Ito, K., Yumoto, N. and Kawata, Y. (1992) *J. Biol. Chem.* **267**, 17773-17779

Morioka, M. and Ishikawa, H. (1993) *J. Biochem.* **114**, 468-472

Mummert, E., Grimm, R., Speth, V., Eckerskorn, C., Schiltz, E., Gatenby, A.A. and Schäfer, E. (1993) *Nature* **363**, 644-648

Nadeau, K., Das, A. and Walsh, C.T. (1993) *J. Biol. Chem.* **268**, 1479-1487

Neupert, W. and Pfanner, N. (1993) *Phil. Trans. R. Soc. Lond. B* **339**, 355-362

Nguyen, V.T., Morange, M. and Bensaude, O. (1989) *J. Biol. Chem.* **264**, 10487-10492

Nigam, S.K., Goldberg, A.L., Ho, S., Rohde, M.F., Bush, K.T. and Sherman, M.Y. (1994) *J. Biol. Chem.* **269**, 1744-1749

Noiva, R., Kimura, H., Roos, J. and Lennarz, W.J. (1991) *J. Biol. Chem.* **266**, 19645-19649

Noiva, R., Freedman, R.B. and Lennarz, W.J. (1993) *J. Biol. Chem.* **268**, 19210-19217

Nozaki, Y. (1972) *Methods Enzymol.* **26**, 43-50

Ostermann, J., Horwich, A.L., Neupert, W. and Hartl, F.-U. (1989) *Nature* **341**, 125-130

Pace, C.N. (1986) *Methods Enzymol.* **131**, 266-280

Palleros, D.R., Welch, W.J. and Fink, A.L. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5719-5723

Palleros, D.R., Reid, K.L., Shi, L., Welch, W.J. and Fink, A.L. (1993) *Nature* **365**, 664-666

Parr, G.R. and Hammes G.G. (1975) *Biochemistry* **14**, 1600-1605

Pelham, H.R.B. (1986) *Cell* **46**, 959-961

Peters, T. Jr. and Davidson, L.K. (1982) *J. Biol. Chem.* **257**, 8847-8853

Picard, D., Khursheed, B., Garabedian, M.J., Fortin, M.G., Lindquist, S. and Yamamoto, K.R. (1990) *Nature* **348**, 166-168

Pihlajaniemi, T., Helaakoski, T., Tasanen, K., Myllylä, R., Huhtala, M.-L., Koivu, J. and Kivirikko, K.I. (1987) *EMBO J.* **6**, 643-649

Price, N.C. (1979) *Biochem. J.* **177**, 603-612

Price, N.C., Kelly, S.M., Wood, S. and auf der Mauer A. (1991) *FEBS Lett.* **292**, 9-12

Price, N.C. (1992) In *Receptor Subunits and Complexes* (Eds. Barnard, E.A., Burgen, A.S.V. and Roberts, G.C.K.) pp 9-38, Cambridge University Press, Cambridge

Price, N.C., Kelly, S.M., Thomson, G.J., Coggins, J.R., Wood, S. and auf der Mauer, A. (1993) *Biochim. Biophys. Acta* **1161**, 52-58

Provencher, S.W. and Glöckner, J. (1981) *Biochemistry* **20**, 33-37

Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zerovnik, E. and Razgulyaev, O.I. (1990) *FEBS Lett.* **262**, 20-24

Rabinkov, A.G. and Amontov, S.V. (1990) *Biochim. Biophys. Acta* **1037**, 216-220

Richarme, G. and Kohiyama, M. (1993) *J. Biol. Chem.* **268**, 24074-24077

Rippmann, F., Taylor, W.R., Rothbard, J.B. and Green, N.M. (1991) *EMBO J.* **10**, 1053-1059

Rospert, S., Glick, B.S., Jenö, P., Schatz, G., Todd, M.J., Lorimer, G.H. and Viitanen, P.V. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10967-10971

Roth, R.A. and Pierce, S.B. (1987) *Biochemistry* **26**, 4179-4182

Roy, S. and Colman, R.F. (1980) *J. Biol. Chem.* **255**, 7517-7520

Rudolph, R., Bohrer, M. and Fischer, S. (1983) *Eur. J. Biochem.* **131**, 383-386

Saibil, H., Dong, Z., Wood, S. and auf der Mauer, A. (1991) *Nature* **353**, 25-26

Sanchez, E.R., Meshinchi, S., Tienrungro, W., Schlesinger, M.J., Toft, D.O. and Pratt, W.B. (1987) *J. Biol. Chem.* **262**, 6986-6891

Sanders, B.M., Nguyen, J., Douglass, T.G. and Miller, S. (1994) *Biochem. J.* **297**, 21-25

Saraste, M., Sibbald, P.R. and Wittinghofer, A. (1990) *Trends Biochem. Sci.* **15**, 430-434

Scherer, P.E., Krieg, U.C., Hwang, S.T., Vestweber, D. and Schatz, G. (1990) *EMBO J.* **9**, 4315-4322

Schmidt, M. and Buchner, J. (1992) *J. Biol. Chem.* **267**, 16829-16833

Schneuwly, S., Shortridge, R.D., Larrivee, D.C., Ono, T., Ozaki, M. and Pak, W.L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5390-5394

Schröder, H., Langer, T., Hartl, F.-U. and Bukau, B. (1993) *EMBO J.* **12**, 4137-4144

Sell, S.M., Eisen, C., Ang, D., Zylicz, M. and Georgopoulos, C. (1991) *J. Bacteriol.* **172**, 4827-4835

Semisotnov, G.V., Rodionova, N.A., Kutyshenko, V.P., Ebert, B., Blank, J. and Ptitsyn, O.B. (1987) *FEBS Lett.* **224**, 9-13

Sherman, M.Y. and Goldberg, A.L. (1992) *Nature* **357**, 167-169

Shieh, B.-H., Stammes, M.A., Seavello, S., Harris, G.L. and Zuker, C.S. (1989) *Nature* **338**, 67-70

Shinde, U., Li, Y., Chatterjee, S. and Inouye, M.(1993) *Proc. Natl. Acad. Sci. USA* **90**, 6924-6928

Siekierka, J.J., Hung, S.H.Y., Poe, M., Lin, C.S. and Sigal, N.H. (1989) *Nature* **341**, 755-757

Silen , J.L. and Agard, D.A. (1989) *Nature* **341**, 462-464

Silver, P.A. and Way, J.C. (1993) *Cell* **74**, 5-6

Skowyra, D, Georgopoulos, C and Zylicz, M. (1990) *Cell* **62**, 939-944

Spangfort, M.D., Surin, B.P., Oppentocht, J.E., Weibull, C., Carlemalm, E., Dixon, N.E. and Svensson, L.A.(1993) *FEBS Lett.* **320**, 160-164

Srivastava, S.K., Abraham, K.I. and Modak, M.J. (1983) *Biochim. Biophys. Acta* **745**, 194-201

Svensson, L.A., Surin, B.P., Dixon, N.E. and Spangfort, M.D. (1994) *J. Mol. Biol.* **235**, 47-52

Takahashi, N., Hayano, T. and Suzuki, M. (1989) *Nature* **337**, 473-475

Tashiro, R., Inoue, T. and Shimozawa, R. (1982) *Biochim. Biophys. Acta* **706**, 129-135

Thomson, G.J., Coggins, J.R. and Price, N.C. (1993) *FEBS Lett.* **336**, 19-22

Tilly, K. and Georgopoulos, C. (1982) *J. Bacteriol.* **149**, 1082-1088

Todd, M.J., Viitanen, P.V. and Lorimer, G.H. (1993) *Biochemistry* **32**, 8560-8567

Trent, J.D., Nimmesgern, E., Wall, J.S., Hartl, F.-U. and Horwich, A.L. (1991) *Nature* **354**, 490-493

Tsalkova, T., Zardeneta, G., Kudlicki, W., Kramer, G., Horowitz, P.M. and Hardesty, B. (1993) *Biochemistry* **32**, 3377-3380

Tseugeki, R., Mori, H. and Nishimura, M. (1992) *Eur. J. Biochem.* **209**, 453-458

Tsou, C.L. (1986) *Trends Biochem. Sci.* **11**, 427-429

Tsou, C.-L.(1988) *Biochemistry* **27**, 1809-1812

van der Vies, S.M., Viitanen, P.V., Gatenby, A.A., Lorimer, G.H. and Jaenicke, R. (1992) *Biochemistry* **31**, 3635-3644

van der Vies, S.M., Gatenby, A.A. and Georgopoulos, C. (1994) *Nature* **368**, 654-656

Viitanen, P.V., Lubben, T.H., Reed, J., Goloubinoff, P., O'Keefe, D.P. and Lorimer, G.H. (1990) *Biochemistry* **29**, 5665-5671

Viitanen, P.V., Donaldson, G.K., Lorimer, G.H., Lubben, T.H. and Gatenby, A.A. (1991) *Biochemistry* **30**, 9716-9723

Viitanen, P.V., Gatenby, A.A. and Lorimer, G.H. (1992a) *Protein Sci.* **1**, 363-369

Viitanen, P.V., Lorimer, G.H., Seetharam, R., Gupta, R.S., Oppenheim, J., Thomas, J.O. and Cowan, N.J. (1992b) *J. Biol. Chem.* **267**, 695-698

Walker, J.E., Saraste, M., Runswick, M.J.and Gay, N.J. (1982) *EMBO J.* **1**, 945-951

Wang, C.-C. and Tsou, C.-L. (1993) *FASEB J.* **7**, 1515-1517

Wang, T.-F., Chang, J. and Wang C. (1993) *J. Biol. Chem.* **268**, 26049-26051

West, S.M. and Price, N.C. (1988) *Biochem. J.* **251**, 135-139

West, S.M. and Price, N.C. (1989) *Biochem. J.* **261**, 189-196

West, S.M. and Price, N.C. (1990) *Biochem. J.* **265**, 45-50

West, S.M., Kelly, S.M. and Price, N.C. (1990) *Biochim. Biophys. Acta* **1037**, 332-336

Wetterau, J.R., Combs, K.A., Spinner, S.N. and Joiner, B.J. (1990) *J. Biol. Chem.* **265**, 9800-9807

Wiech, H., Buchner, J., Zimmermann, R. and Jakob, U. (1992) *Nature* **358**, 169-170

Winther, J.R. and Sørensen, P. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9330-9334

Wynn, R.M., Davie, J.R., Cox, R.P. and Chuang, D.T. (1992) *J. Biol. Chem.* **267**, 12400-12403

Yaffe, M.B., Farr, G.W., Miklos, D., Horwich, A.L., Sternlicht, M.L. and Sternlicht, H. (1992) *Nature* **358**, 245-248

Zheng, X., Rosenberg, L.E., Kalousek, F. and Fenton, W.A. (1993) *J. Biol. Chem.* **268**, 7489-7493

Zhu, X., Ohta, Y., Jordan, F. and Inouye, M. (1989) *Nature* **339**, 483-484

Zweig M. and Cummings, D.J. (1973) *J. Mol. Biol.* **80**, 505-518

Zylicz, M. (1993) *Phil. Trans. R. Soc. Lond. B* **339**, 271-278

APPENDIX

The following publications contain some of the results obtained in the present work

1. Thomson, G.J., Coggins, J.R. and Price, N.C. (1992) *Biochem. Soc. Trans.* **21**, 61S

2. Price, N.C., Kelly, S.M., Thomson, G.J., Coggins, J.R., Wood, S. and auf der Mauer, A. (1993) *Biochim. Biophys. Acta* **1161**, 52-58

3. Thomson, G.J., Coggins, J.R. and Price, N.C. (1993) *FEBS Lett.* **336**, 19-22