a milk bottle for 10 or 12 min. The plate was then placed in a bottle containing yeasted cornmeal; emerging adults were scored and selected 10 days later.
6. B. Charlesworth and D. Charlesworth, *Heredity* **54**, 71 (1985). The Ives strain was initially isolated in Massachusetts more than 20 years ago and was put through a bottleneck of 50 pair matings, which removed any segregating, cytologically visible inversions.
7. The first two experiments at Stanford were performed with flies obtained from G. Spicer. These flies were established from B. Charlesworth's stock, which was used for the later repetition at Michigan. The stocks used have been isolated for at least 4 years, and both have been maintained by ordered rotation of 10 bottles each generation.
8. The heritability (h^2) of phenocopy liability can be estimated by following Falconer's method for threshold-dependent traits [D. S. Falconer, *Introduction to Quantitative Genetics* (Longman, New York, ed. 2, 1981), p. 272]: $h^2 = (x_P - x_R)/i$ where x_P and x_R are the proportions of phenocopied individuals in the parental (Ives) and progeny populations, respectively, and i is the intensity of selection, all expressed in standard deviation units. For percentages of 13 and 24% (the mean frequencies in the F_1 generation of the three selection experiments) the values were as follows: $x_P = 1.126$, $x_R = 0.706$, $i = 1.627$, and $h^2 = 0.26$. However, the error associated with the measurements was high: The ratio of genetic to phenotypic variance was in the range of 0.15 to 0.40, if an underlying normal distribution of threshold-dependent liability is assumed.
9. E. B. Lewis, *Nature* **276**, 565 (1978); I. Duncan, *Annu. Rev. Genet.* **21**, 285 (1987).
10. M. P. Capdevila and A. García-Bellido, *Wilhelm Roux's Arch. Dev. Biol.* **190**, 339 (1981).
11. Five of 17 sets of discs from larvae treated after six generations of up-selection showed patches of lost expression. This finding is in agreement with the frequency of phenocopies seen in these animals. Zero of 12 sets of discs from animals not treated with ether showed loss of UBX expression.
12. E. P. H. Yap and J. O. McGee, *Trends Genet.* **8**, 49 (1992).
13. In addition to the four Bithorax complex polymorphisms discussed in the text, two sequences sampled from the coding region appeared to be monomorphic, one sequence from the *bxd* region amplified inconsistently, and a fifth polymorphism in the *pbx* region within 5 kb of A/a gave genotype values almost identical to A/a.
14. The primer locations are contained in the complete sequence of the Bithorax complex (GenBank accession number U31961). Primer pairs for PCR amplification were as follows: A/a, ACATGAAAACATTTGCGTAA and TTATGCGCCGTCGCTCTAAA; C/c, ACGTTGCATTTGCCGTTGCA and AGAAGGGGTGGTGCA (originally from unpublished *Ubx* sequence, D. Peattie and D. S. Hogness); D/d, TTGAGAGAGTCTTCGCG and GGTATCGGTATGGTATCGG [nt 3226 to 3362 of (*17*)]; and Z/z, GCACCACCCACAGGTGCA and TTGCTCGCATTCAACATT [nt 257 to 745 of (*18*)]. PCR reactions were carried out in an Ericomp or Perkin-Elmer machine, with 40 cycles. Each cycle consisted of 30 min at 94°C, 30 min at 57°C, and 45 min at 72°C and was started with about 50 ng of genomic DNA prepared from a single fly, in 25-μl reactions at 1 mM Mg. For alleles A and C, 10 μl of product was run directly into a 6.5% polyacrylamide gel in 1 × TAE (0.04 M tris and 1 mM EDTA adjusted to pH 7.4 with acetate) maintained at 25°C; polymorphisms were detected as heteroduplexes. Samples of alleles D and Z (9 μl) were first denatured with 1 μl of 0.1 M NaOH and 5 mM EDTA for 5 min at 37°C and then with 1 μl of formamide before they were loaded. Samples of allele Z were also digested before denaturation with Pvu II restriction enzyme to create products small enough to resolve by SSCP (*12*). Bands were detected by ethidium bromide staining or by radiography after inclusion of a trace amount of ^{35}S-labeled dATP in the PCR reactions.
15. K. D. Irvine, S. Helfand, D. S. Hogness, *Development* **111**, 407 (1991); J. Simon, M. Peifer, W. Bender, M. O'Connor, *EMBO J.* **9**, 3945 (1990).
16. A. Busturia, J. Casanova, E. Sanchez-Herrero, R. Gonzales, G. Morata, *Development* **107**, 575 (1989).
17. K. Kornfeld *et al.*, *Genes Dev.* **3**, 243 (1989).
18. F. Karch, W. Bender, B. Weiffenbach, *ibid.* **4**, 1573 (1990).
19. C. Zapata and G. Alvarez, *Evolution* **46**, 1900 (1992).
20. J. N. Macpherson, B. S. Weir, A. J. Leigh Brown, *Genetics* **126**, 121 (1990).
21. M. E. Akam, H. Moore, A. Cox, *Nature* **309**, 635 (1983).
22. C. F. Aquadro *et al.*, *Genetics* **132**, 443 (1992); W. G. Hill and A. Robertson, *Theor. Appl. Genet.* **38**, 226 (1968).
23. *GS1*: R. Caizzi, M. P. Bozzetti, C. Caggese, F. Ritossa, *J. Mol. Biol.* **212**, 17 (1990); *EcR*: M. R. Koelle *et al.*, *Cell* **67**, 59 (1991); *bicoid*: T. Berleth *et al.*, *EMBO J.* **7**, 1749 (1988).
24. The contribution of the major *Ubx* polymorphism to the genetic variance can be estimated from the relation $h^2 = 2pq[a + (q - p)d]/s^2$ where a and d are the additive and dominance effects associated with the polymorphism. Estimates based on determination of the mean frequencies of 15 isofemale lines that are homozygous or heterozygous for either the A or a alleles yield values of $a = 11.8$ and $d = -6.8$ (G. Gibson, unpublished data) against a total standard deviation of 16.0 percentage units. For $p = 0.39$ in the Ives strain, h^2 due to *Ubx* was 0.196. The arcsin ($\sqrt{x}$) transformed data yield an estimate of 0.170. If the departure from additivity due to recessivity is ignored, this estimate suggests that a polymorphism linked to *Ubx*-A would contribute between 65 and 75% of the genetic variance, given the overall heritability estimate in (*8*). However, given the error associated with all measurements and uncertainties about the additivity or normality, or both, of the underlying liability distribution, the estimate has very low confidence. Better estimates await identification of the remaining liability-effect loci.
25. Resistant (or sensitive) virgin females were crossed en masse to sensitive (or resistant) males, and phenocopy frequencies were determined from multiple trials. Significant results were obtained with stocks at the conclusion of both the second Stanford experiment (20% compared with 5%, 10-min ether treatments; $P < 0.005$, $n = 5$ trials, t test) and the Michigan experiment [35% compared with 25%, 15-min ether treatments; $P < 0.02$, $n = 6$; the parents in this case were from isofemale lines and showed extreme sensitivity (55%) and resistance (2%), respectively]. Comparison of males and females independently yielded identical conclusions, indicating that there was no strong contribution from the X chromosome.
26. C. Lai, R. F. Lyman, A. D. Long, C. H. Langley, T. F. C. Mackay, *Science* **266**, 1697 (1994); T. F. C. Mackay and C. H. Langley, *Nature* **348**, 64 (1990); S. Tanksley, *Annu. Rev. Genet.* **27**, 205 (1993).
27. L. A. Zhivotovsky and S. Gavrilets, *Theor. Popul. Biol.* **42**, 254 (1992).
28. We thank B. Charlesworth and G. Spicer for supplying us with the Ives strains, M. Bender for assistance in the early stages, and R. Fuller and C. Goodnow for the use of their PCR machines. G.G. particularly thanks W. Watt, M. Feldman, and J. Adams for their support and advice and D. Pollock, M. Montano, D. Andrew, and members of the Feldman and Hogness laboratories for comments and discussions. Supported by grants from the Rackham Graduate School of the University of Michigan and by grants to D.S.H. from the National Institutes of Health. G.G. was supported by a postdoctoral fellowship from the Helen Hay Whitney Foundation.

14 August 1995; accepted 24 October 1995

Structure of the Heat Shock Protein Chaperonin-10 of *Mycobacterium leprae*

Shekhar C. Mande, Vijay Mehra, Barry R. Bloom, Wim G. J. Hol*

Members of the chaperonin-10 (cpn10) protein family, also called heat shock protein 10 and in *Escherichia coli* GroES, play an important role in ensuring the proper folding of many proteins. The crystal structure of the *Mycobacterium leprae* cpn10 (Ml-cpn10) oligomer has been elucidated at a resolution of 3.5 angstroms. The architecture of the Ml-cpn10 heptamer resembles a dome with an oculus in its roof. The inner surface of the dome is hydrophilic and highly charged. A flexible region, known to interact with cpn60, extends from the lower rim of the dome. With the structure of a cpn10 heptamer now revealed and the structure of the *E. coli* GroEL previously known, models of cpn10:cpn60 and GroEL:GroES complexes are proposed.

Mycobacteria are among the most important human microbial pathogens. *Mycobacterium tuberculosis* is estimated to be responsible for 2.02 million deaths per year, particularly in developing countries, and has recently reemerged in the industrialized countries (*1–3*). *Mycobacterium leprae* causes a disfiguring disease, leprosy, in 30% of the untreated cases. Mycobacteria have many unusual features, one of the most remarkable being the ability of these organisms to reside in Schwann cells and macrophages, the latter being the very cells that should destroy the invading organisms. One of the important antigens of *M. leprae* recognized by T cells is cpn10, a 10-kD heat shock protein. In patients with tuberculoid leprosy, approximately one-third of the *M. leprae*–reactive T cells that respond to the whole organism also respond to cpn10 (*4*). We report here the crystal structure of *M. leprae* cpn10 (Ml-cpn10).

Members of the cpn10 family, GroES in

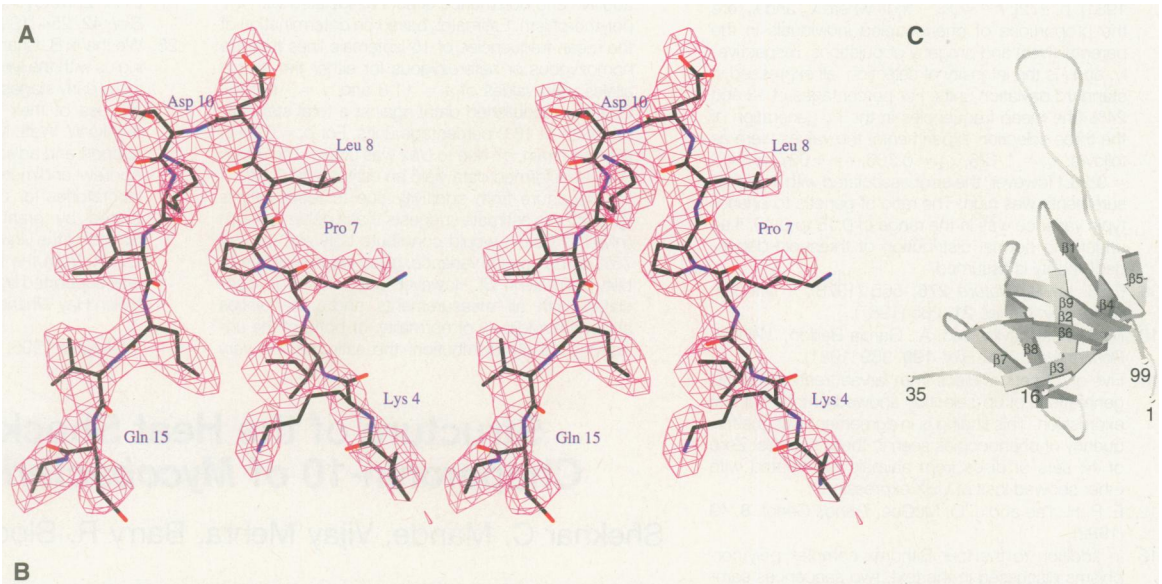
S. C. Mande, Department of Biological Structure and Biomolecular Structure Center, University of Washington, Seattle, WA 98195, USA.
V. Mehra, Department of Microbiology and Immunology, University of Washington, Seattle, WA 98195, USA.
B. R. Bloom, Department of Microbiology and Immunology, University of Washington, Seattle, WA 98195, USA, and Howard Hughes Medical Institute, Albert Einstein College of Medicine, Bronx, NY 10461, USA.
W. G. J. Hol, Department of Biological Structure and Biomolecular Structure Center, and Howard Hughes Medical Institute, University of Washington, Box 357742, Seattle, WA 98195, USA. E-mail: hol@xray.bmsc.washington.edu

*To whom correspondence should be addressed.

Escherichia coli, are known to interact with members of the cpn60 family, GroEL in *E. coli*, forming a complex important for the proper folding of several proteins (*5–7*). Cpn60 is a 14-unit cylindrical protein that consists of two rings of seven protomers forming a hollow inner cage and two wide open cavities at each end (*8*). Cpn10 and cpn60 can form two different complexes. In one, the cpn60 cylinder is capped on both sides by a cpn10 heptamer, yielding a $cpn60_{14}$:$cpn10_{14}$ symmetric complex (*9*). The second type of complex is asymmetric with only one cpn10 heptamer bound to a cpn60 tetradecamer (*10*). The bound cpn10 increases the adenosine diphosphate affinity of cpn60 by about four orders of magnitude (*6*) and, in a process for which many aspects still need to be unraveled, allows the cpn60 tetradecamer to use adenosine triphosphate (ATP) in the refolding of the substrate protein (*6, 7, 11*).

The crystal structure of Ml-cpn10 was solved by the isomorphous replacement method with a single poor derivative that yielded phases only up to ~5 Å resolution (*12*). In the resulting single isomorphous replacement map, the outline of the molecule was quite clear. Subsequent solvent flattening and sevenfold density averaging allowed a dramatic improvement of the electron density distribution, which could be correlated unambiguously with the amino acid sequence of Ml-cpn10 (Fig. 1, A and B). Only one region of the molecule, comprising residues 17 to 35, has discontinuous density, and this is exactly the region in *E. coli* GroES that has been reported to be flexible (*13*).

The 99-residue polypeptide chain of a single Ml-cpn10 subunit has a triangular shape with dimensions of approximately 37 Å by 32 Å by 37 Å and with a distinct hydrophobic



B

```
                1        10        20        30        40        50        60        70        80        90        99
Mycle  .........A KVKIKPLEDK ILVQAGEAET MTPSGLVIPE NAKEKPQEGT VVAVGPGRWD EDGAKRIPVD VSEGDIVIYS KYGGTE.IKY NGEEYLILSA RDVLAVVSK. .
Mycbo  .........A KVNIKPLEDK ILVQANEAET TTASGLVIPD TAKEKPQEGT VVAVGPGRWD EDGEKRIPLD VAEGDTVIYS KYGGTE.IKY NGEEYLILSA RDVVGRRFK. ..
Myctu  ........MA KVNIKPLEDK ILVQANEAET TTASGLVIPD TAKEKPQEGT VVAVGPGRWD EDGEKRIPLD VAEGDTVIYS KYGGTE.IKY NGEEYLILSA RDVLAVVSK. ..
Human  ......AGQA FRKFLPLFDR VLVERSAAET VTKGGIMLPE KSQGKVLQAT VVAVGSGSKG KGG.EIQPVS VKVGDKVLLP EYGGTKVVLD .DKDYFLFRD GDILGKYVD. ..
  Rat  ......AGQA FRKFLPLFDR VLVERSAAET VTKGGIMLPE KSQGKVLQAT VVAVGSGGKG KGG.EIQPVS VKVGDKVLLP EYGGTKVVLD .DKDYFLFRD GDILGKYVD. ..
Arath  .........M MKRLIPTFNR ILVQRVIQPA KTESGILLPE KS.SKLNSGK VIAVGPGSRD KDG.KLIPVS VKEGDTVLLP EYGGTQVKLG .ENEYHLFRD EDVLGTLHED ..
Yeast  ...MSTLLKS AKSIVPLMDR VLVQRIKAQA KTASGLYLPE KNVEKLNQAE VVAVGPGFTD ANG.NKVVPQ VKVGDQVLIP QFGGSTIKLG NDDEVILFRD AEILAKIAKD ..
Acyps  .......... .MKIRPLHDR VLVKRQEVES KSAGGIVLTG SAAGKSTRGT VTAVGKGRVL DNG.DIKPLD VKVGDVVIPN EGYGAKTEKI DNEELLILTE SDILAIVE.. ..
Haedu  .......... .MSIRPLHDR VILKREEVET CSAGGIVLTG SATVKSTRGK VIAVGTGRLL ENG.SIQPLS VKEGDMVIFN EGYGAKVEKI DGEEILILSE NDILAIVE.. ..
Ecoli  .......... .MNIRPLHDR VIVKRKEVET KSAGGIVLTG SAAAKSTRGE VLAVGNGRIL ENG.EVKPLD VKVGDIVIFN DGYGVKSEKI DNEEVLIMSE SDILAIVEA. ..
Chlpn  ....MSDQAT TLRIKPLGDR ILVKREEEEA TARGGIILPD TAKKKQDRAE VLVLGTGKRT DDG.TLLPFE VQVGDIILMD KYAGQE.ITI DDEEYVILQS SEIMAVLK.. ..
Chlps  ....MSDQAT TLRIKPLGDR ILVKREEEDS TARGGIILPD TAKKKQDRAE VLVLGTGKRD KDG.NVLPFE VTVGDTVLID KYAGQE.LTV DGEEYVIVQE SEVMAVLK.. ..
Chltr  ....MSDQAT TLKIKPLGDR ILVKREEEAS TARGGIILPD TAKKKQDRAE VLALGTGKKD DKG.QQLPFE VQVGDIVLID KYSGQE.LTV EGEEYVIVQM SEVIAVLQ.. ..
Stral  ....MTTASS KVAIKPLEDR IVVQPLDAEQ TTASGLVIPD TAKEKPQEGV VLAVGPGRFE NGE..RLPLD VKTGDVVLYS KYGGTE.VKY NGEEYLVLSA RDVLAIIEK. ..
Strco  ....MTTTSS KVAIKPLEDR IVVQPLDAEQ TTASGLVIPD TAKEKPQEGV VLAVGPGRFE DGN..RLPLD VSVGDVVLYS KYGGTE.VKY NGEEYLVLSA RDVLAIVEK. ..
Lepin  .......... MASIKPLGDR VLVEPRQEAE EKIGSIFVPD TAKEKPQEGK VVEIGSGKYE DG..KLIPLE VKVGDTVLYG KYSGTE.IKS EGKEYLIIRE SDILAVVKK. ..
Cloab  .......... .MKIRPLGDR VVIKRLEAEE TTKSGIVLPS SAKEKPQMAE VVAVGPGGVV D.G.KEIQMQ VKTGDKVFFS KYSGTE.IKV DNEELLILRQ DDILGIVEE. ..
Clope  .......... .MSIKPLGDR VVIKRLEAEE TTKSGIIVTG TAKERPQEAE VVAVGPGAIV D.G.KRTEME VKIGDKVLYS KYAGTE.VKF EGEEYTILRQ DDILAIVE.. ..
Agrtu  ........MT STNFRPLHDR VVVRRVESEA KTKGGIIIPD TAKEKPQEGE IVAVGSGARD EAG.KVVALD VKVGDRVLFG QWSGTE.VKL DGEDSSIMKE ADIMGIIG.. ..
Bruab  ........MA DIKFRPLHDR VVVRRVESEA KTAGGIIIPD TAKEKPQEGE VVAAGAGARD EAG.KLVPLD VKAGDRVLFG KWSGTE.VKI GGEDLLIMKE SDILGIVG.. ..
Legmi  .......... .MKIRPLHDR VVVRRMEEER TTAGGIVIPD SATEKPTRGE IIAVGPGKVL ENG.DVRALA VKVGDVVLFG KYSGTE.VKI SGQELVVMRE DDIMGVIEK. ..
Legpn  .......... .MKIRPLHDR VVVRRMEEER TTAGGIVIPD SATEKPMRGE IIAVGAGKVL ENG.DVRALA VKVGDVVLFG KYSGTE.VKV DGKELVVMRE DDIMGVIEK. ..
Amops  .......... .MKIRPLHDR VVVRRLEEER TTAGWIVIPD SATEKPMRGE IIAIGAGKIL DNG.DVRAFV VKVGDVVLFG KYSGTE.VKV AGQELVVMRE DDIMGVIEK. ..
Coxbu  .......... .MKIRPLHDR VVVRRLEEER TSAGGIVIPD SAAEKPSRGE VISVGPGKPL DNG.EVRSLD VKVGDQILFG KYAGTE.VKL AGDEYIVMRE DDIMGVIEK. ..
Chrvi  .......... .MNIRPLHDR VVVRRMEEER LSAGGIVIPD SATEKPIQGE IIAVGHGKIL DNG.SVRALD VKVGDSVLFG KYSGTE.VKL DGKEFLVMRE EDIMAVVEG. ..
Pseae  .......... .MKLRPLHDR VVIRRSEEET KTAGGIVLPG SAAEKPNRGE VVAVGTGRVL DNG.EVRALA VKVGDKVVFG PYSGSNAIKV DGEELLVMGE SEILAVLED. ..
Synp6  ...MAAVSLS VSTVTPLGDR VFVKVAEAEE KTAGGIILPD NAKEKPQVGE IVAVGPGKRN DDG.SRQAPE VKIGDKVLYS KYAGTD.IKL GNDDYVLLSE KDILAVVA.. ..
Synp7  ...MAAVSLS VSTVTPLGDR VFVKVAEAEE KTAGGIILPD NAKEKPQVGE IVAVGPGKSN DDG.SRQAPE VKIGDKVLYS KYAGTD.IKL GNDDYVLLSE KDILAVVA.. ..
Syny3  ...MAAISIN VSTVKPLGDR VFVKVSPAEE KTAGGILLPD NAKEKPQIGE VVQVGPGKRN DDG.TYSPVE VKVGDKVLYS KYAGTD.IKL GGDDYVLLTE KDILASVA.. ..
Bacst  .......... ..MLKPLGDR VVIEVIETEE KTASGIVLPD TAKEKPQEGR VVAVGKGRVL DSG.ERVAPE VEVGDRIIFS KYAGTE.VKY DGKEYLILRE SDILAVIG.. ..
Thep3  .......... ..MLKPLGDR IVIEVVETEE KTASGIVLPD TAKEKPQEGR VVAVGAGRVL DNG.QRIGRK SKVGDRVIFS KYAGTE.VKY DGKEYLILRE SDILAVIR.. ..
Bacsu  .......... ..MLKPLGDR VVIELVESEE KTASGIVLPD SAKEKPQEGK IVAAGSGRVL ESG.ERVALE VKEGDRIIFS KYAGTE.VKY EGTEYLILRE SDILAVIG.. ..
Staau  .......... ..MLKPIGNR VIIEKKEQEQ TTKSGIVLTD SAKEKSNEGV IVAVGTGRLL NDG.TRVTPE VKEGDRVVFQ QYAGTE.VKR DNETYLVLNE EDILAVIE.. ..
Lacla  .......... ..MLKPLENR VVLRVKEEEE KSMGGIVLTS ASQEKPQTAE VVAVGEGKTN HHG.TLISPL VKVGDTVIFE KFAGTT.VKM DGEEFLILKD SDLLAIVE.. ..
 Spl2  VRAASITTSK YTSVKPLGDR VLIKTKIVEE KTTSGIFLPT AAQKKPQSGE VVAIGSGKKV ..GDKKLPVA VKTGAEVVYS KYTGTE.IEV DGSSHLIVKE DDIIGILETD DV
 Spl1  .......... .KDLKPLNDR LLIKVAEVEN KTSGGLLLAE SSKEKPSFGT VVATGPGVLD EEG.NRIPLP VCSGNTVLYS KYAGNDFKGV DGSDYMVLRV SDVMAVLS.. ..
Ricts  .......... .MKYQPLYDR VLVEPIQNDE .AHGKILIPD TAKEKPTEGI VVMVGGGYRN DKG.DITPLK VKKGDTIVYT KWAGTE.IKL ESKDYVVIKE SDILLVKS.. ..
```


Fig. 1. Sequence information and course of polypeptide chain of Ml-cpn10. (**A**) Electron density map in stereo showing the conserved "signature" residues 7 to 10 "PLXD" (*30*) of Ml-cpn10. This 3.5 Å–resolution map was calculated with observed amplitudes and phases obtained by starting from the 5 Å single isomorphous replacement map, which was subsequently extended in 150 cycles to 3.5 Å by solvent flattening and symmetry averaging (*12*). The map is therefore completely unbiased by the model. (**B**) Alignment of the known cpn10 sequences as obtained by the Genetic Computer Group (*31*). Antigenic loops of Ml-cpn10, the 10-kD antigen (*15*), are shown in purple. The flexible loop of *E. coli* GroES (*13*) is marked in green. The residues highlighted in blue show a high degree of sequence conservation and are part of the proposed GroEL interaction site. Residues involved in three charged rings lining the inner surface of cpn10 are marked in orange. The segment PLXD highlighted in pink is conserved in all the known sequences and appears to be the signature tetrapeptide of the cpn10 family. (**C**) Conformation of a single subunit of Ml-cpn10. Only the well-defined parts of the molecule (residues 1 to 16 and 35 to 99) are shown. The five-stranded antiparallel β sheet A consists of residues 76 to 80 (β7), 84 to 88 (β8), 12 to 15 (β2), 41 to 44 (β4), and 63 to 65 (β5). The four-stranded antiparallel β sheet B comprises residues 38 to 41 (β3), 67 to 70 (β6), 94 to 98 (β9), and 4′ to 8′ (β1′) of another subunit. Because of the limited resolution of the Ml-cpn10 structure, these visually derived β strand–defining residue numbers are approximate. Residues contributing to the hydrophobic core of the molecule are Ile^{12}, Val^{14}, Val^{42}, Val^{68}, Tyr^{70}, Leu^{88}, and Val^{93}, which are highly conserved in the cpn10 family [see (B)].

core formed by two orthogonal antiparallel β sheets. The five-stranded β sheet A is formed by β strands 7, 8, 2, 4, and 5, whereas the four-stranded sheet B is formed by β strands 3, 6, and 9 plus β strand 1′ provided by a neighboring subunit (Figs. 1C and 2A).

The seven Ml-cpn10 monomers are arranged in a dome-like structure that has close to sevenfold symmetry for almost all residues, except for the flexible loop (Fig. 2A). The interactions between the subunits are mediated mainly by the two antiparallel β strands, β1′ and β9. The hydrophobic interface between the two subunits is formed by residues Val[3], Ile[5], Leu[8], Ile[78], and Ile[88] of one monomer, and Pro[30], Ile[69], Val[93], Leu[94], Ala[95], Val[96], and Val[97] of another monomer (Fig. 2B). All of these residues except Val[3] and Val[97] are conserved as apolar residues in the known cpn10 sequences (Fig. 1B). Each monomer buries 1765 Å^2, constituting as much as 28% of its total surface area, on heptamer formation. Of the total buried area, 1291 Å^2, or 73%, is apolar. The diameter of the heptamer is ~80 Å perpendicular to the sevenfold axis and ~35 Å measured parallel to the axis (Fig. 3A). The top of the dome has an opening with a cross section of ~12 Å. The architecture is strikingly similar to that of the Pantheon in Rome.

The inner surface of the dome is completely hydrophilic with a total of 63 positively and negatively charged residues arranged in concentric rings. The oculus of the dome is lined by a ring of negative carboxylates provided by Asp[51], Glu[52], and Asp[53] of each subunit. Subsequent rings are formed by the positive charges of Arg[49] and

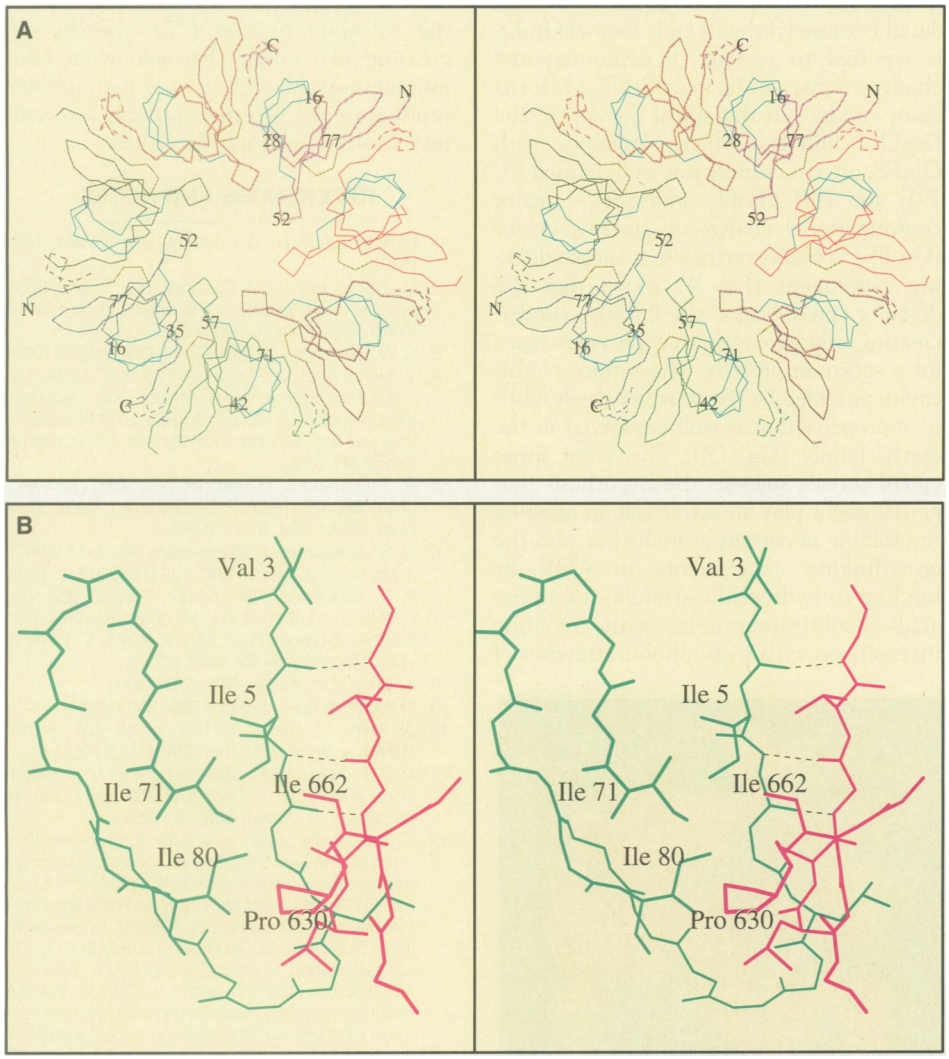


Fig. 2. Structure of the Ml-cpn10 heptamer. (**A**) Stereo view of the Cα trace of the heptameric molecule with each subunit in a different color and key features color-coded. The two major T cell epitopes (*15*) of the molecule are shown in cyan. One of these encompasses the loop of residues 48 to 63 forming the "finger" extending to the oculus in the center of the heptamer. Residues shown in dashed lines could not be assigned an unambiguous sequence number but must form part of residues 17 to 34, known to be flexible in GroES (*13*). The dashed residues of this flexible loop, shown in between each pair of monomers in the heptamer, are most likely connected to the subunits of the crystallographically related heptamer, as depicted in Fig. 3. (**B**) The monomer-monomer interface. The two monomers are shown in blue and pink. The first strand of one monomer and the last strand of the second monomer form an antiparallel β sheet. The interface is predominantly formed by hydrophobic residues, which virtually all are conserved as apolar residues in the entire cpn10 family (see Fig. 1B and text).

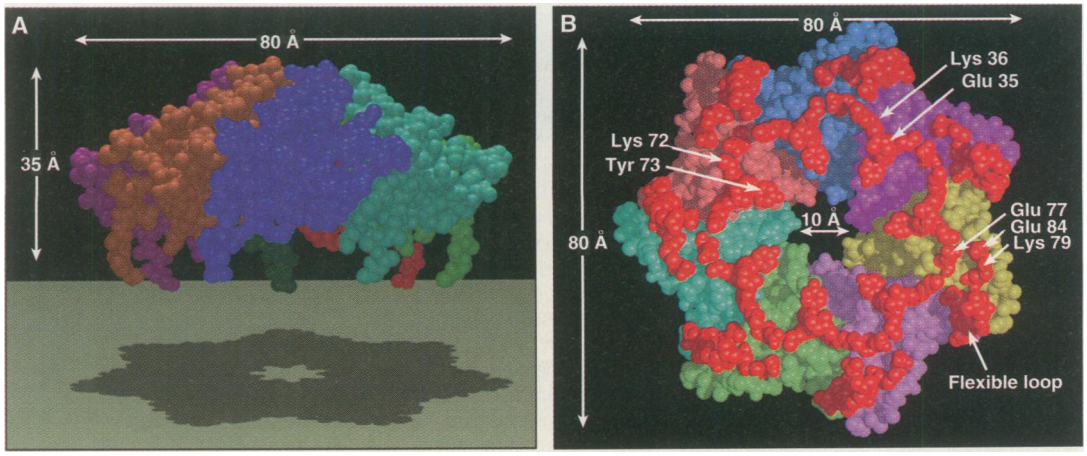


Fig. 3. Characteristic features of the Ml-cpn10 heptameric molecule (*29*). (**A**) Side view of the "dome" perpendicular to the sevenfold axis. The central opening, or "oculus," is seen in the shadow of the heptamer. The flexible loops are assumed to extend downward from the dome. In the crystals, two Ml-cpn10 heptamers interact with each other by juxtaposing their rims, forming a heptamer dimer with approximate D_7 symmetry. The sevenfold axes of the upper and lower heptamers make an angle of ~7° with each other. The distance between the upper and lower oculi of this tetradecamer of Ml-cpn10 is ~72 Å. (**B**) Conserved features of the heptamer, viewed along the sevenfold axis. The flexible loop and other potential residues for GroEL interactions are shown in red. These features would be "seen" by members of the cpn60 family, such as GroEL, when docking with the cpn10 dome.

Arg^{91}, the negative charges of Glu^{9} and Asp^{10}, and, near the rim of the dome, by Lys^{11} and Asp^{92}. The lower rim of the dome contains several well-conserved residues such as Glu^{35}, Lys^{36}, Lys^{72}, and Tyr^{73}, and of particular interest is the flexible loop, which probably extends from the rim "downward," ready to interact with the cpn60 tetradecamer (Fig. 3B). Lys^{36}, which has been implicated to modulate the allosteric transitions in GroEL (*14*), is fully solvent-exposed and appears to be well positioned to interact with GroEL.

Residues 25 to 42 and 57 to 71 have been reported to be the immunodominant T cell epitopes of Ml-cpn10 (*15*). One of these epitopes appears to compose a part of the flexible loop. The other contains part of the long loop formed by residues 48 to 63, which reaches out from the "body" of each subunit to the edge of the oculus (Fig. 2A). The two epitopes are spatially close together.

The three-dimensional structure of Ml-cpn10 reported here can be combined with the structure of *E. coli* GroEL (*8*) to form a model of a cpn60:cpn10 assembly, guided by numerous electron microscopy studies on GroEL, GroES, and their complexes (*9*, *10*, *16*). In our model the source of the GroEL structure is *E. coli*, whereas the source of the GroES structure is *M. leprae*. Because (i) the cpn60's and cpn10's from different sources have been shown to form hybrid complexes and aid in the folding of some proteins (*17*), and (ii) circular dichroism studies reveal structural similarity among cpn10 proteins from different species (*18*), our model can provide insight into the overall architecture of GroEL:GroES complexes. We refrain from going into great detail because (i) the flexible loop of GroES is reported to undergo a conformational change on interacting with GroEL (*13*); (ii) many residues in the apical domain of the GroEL cylinder, which interacts with GroES, are structurally not well defined (*8*, *19*); and (iii) GroEL undergoes a major conformational change on binding cpn10 (GroES) (*16*). Nevertheless, a simple docking experiment (Fig. 4) shows that the dome of cpn10 "caps" the GroEL cylinder, creating a large cavity that provides space for a substrate protein. The surface of this cavity provided by cpn10 is very hydrophilic, a property that is well conserved in the cpn10 family (Fig. 1B). This polar inner cpn10 surface suggests the hypothesis that cpn10 might play an active role in assisting the folding of certain proteins [see also the cross-linking experiments in (*20*)] by binding to hydrophilic stretches of amino acids of substrate proteins, while avoiding interactions with hydrophobic stretches of amino acids, which are to form eventually the interior of the substrate protein. The importance of the GroEL:GroES, or "cis," cavity in the asymmetric complex for productive release of folded protein has recently been emphasized (*21*). The function of GroEL, on the other hand, might be the formation of temporary, ATP-driven interactions with hydrophobic parts of the substrate protein (*22*), thereby decreasing free energy barriers when rearranging misfolded clusters of hydrophobic residues of the substrate protein into correctly folded hydrophobic cores.

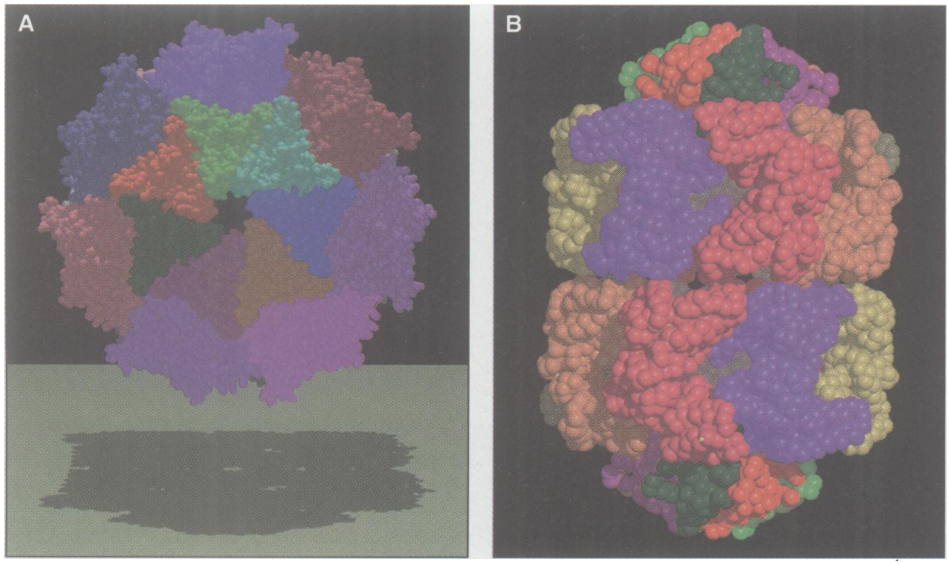


Fig. 4. Models of GroEL:GroES (cpn60:cpn10) complexes based on the structures of Ml-cpn10 (Figs. 2 and 3) and *E. coli* GroEL (*8*). The sevenfold axes of Ml-cpn10 and GroEL are aligned and rotated such that the flexible loops of Ml-cpn10 fit into the "crevices" of GroEL between the apical domains. The complexes depicted show the cpn60:cpn10 complexes just before the apical domain of GroEL undergoes a conformational change (*16*) and are representative for the entire family because of the high sequence conservation among the cpn60 (*5*, *32*) and cpn10 proteins (Fig. 1B). (**A**) Ml-cpn10 structure docked onto the structure of GroEL (*8*) viewed along the sevenfold axis. The outer diameter of the GroEL cylinder is ~135 Å, that of the Ml-cpn10 (GroES) dome is ~80 Å. The asymmetric structure shows that the dimensions of GroES would be sufficiently large to block the entry, and escape, of substrate proteins into the GroEL central cavity on one side. (**B**) Side view of an approximate model of the symmetric complex, which might have a transient function in the chaperonin cycle showing the "(American) football" shape of the tetradecameric GroEL associated with two heptamers of Ml-cpn10. The "windows" of GroEL (*8*) allow a view in the central cavity of the cylinder, whereas the Ml-cpn10 (GroES) domes virtually completely close off the two cavities at the ends of the symmetric GroEL cylinder.

REFERENCES AND NOTES

1. B. R. Bloom and C. J. L. Murray, *Science* **257**, 1055 (1992).
2. C. J. L. Murray and A. D. Lopez, Eds., *Global Comparative Assessments in the Health Sector* (World Health Organization, Geneva, 1994), p. 54.
3. S. W. Hunter *et al.*, *J. Biol. Chem.* **265**, 14065 (1990).
4. V. Mehra *et al.*, *J. Exp. Med.* **175**, 275 (1992). *Mycobacterium leprae* cpn10 shares 90% sequence identity with the corresponding cpn10 of *M. tuberculosis* and 44% sequence identity with the chaperonin GroES of *E. coli*.
5. R. J. Ellis and S. M. van der Vies, *Annu. Rev. Biochem.* **60**, 321 (1991); F. U. Hartl and J. Martin, *Curr. Opin. Struct. Biol.* **5**, 92 (1995).
6. G. S. Jackson *et al.*, *Biochemistry* **32**, 2554 (1993).
7. J. Martin *et al.*, *Nature* **366**, 228 (1993); M. J. Todd, P. V. Viitanen, G. H. Lorimer, *Science* **265**, 659 (1994); S. J. Landry and L. M. Gierasch, *Annu. Rev. Biophys. Biomol. Struct.* **23**, 645 (1994); T. Tsalkova *et al.*, *Biochemistry* **32**, 3377 (1993).
8. K. Braig *et al.*, *Nature* **371**, 578 (1994).
9. M. Schmidt *et al.*, *Science* **265**, 656 (1994).
10. A. Azem, M. Kessel, P. Goloubinoff, *ibid.*, p. 653 (1994); A. Engel *et al.*, *ibid.* **269**, 832 (1995).
11. M. J. Todd *et al.*, *ibid.* **265**, 659 (1994); M. K. Mayer, F. U. Hartl, J. Martin, F. U. Hartl, *ibid.* **269**, 836 (1995); H. R. Saibil *et al.*, *Nature* **353**, 25 (1991).
12. Crystals were grown at 4°C by the hanging drop method. The best crystals were obtained by mixing 3 μl of protein solution in 5 mM MES buffer (pH 6.0), 1 mM EDTA, and 1 mM NaN_3 with an equal amount of the well solution containing 100 mM Hepes buffer (pH 7.5), 35% v/v PEG400, and 150 mM Li_2SO_4. The crystals belong to space group $C222_1$ with a = 112.9 Å, b = 129.0 Å, and c = 109.5 Å. Extreme radiation sensitivity made cryo-cooling imperative. A major obstacle was the variation in cell dimensions from crystal to crystal by as much as 6 Å. Eventually seven different data sets from crystals with similar cell dimensions could be merged with an overall merging R factor at 3.5 Å resolution of 0.088 and in the last resolution shell of 0.15. The data are 98% complete to 3.5 Å resolution and 90% complete in the 3.6 to 3.5 Å resolution range. Self-rotation functions calculated showed a strong peak for a rotation of 51.4° in a direction virtually parallel to the crystal a-direction, thereby identifying the molecular sevenfold axis. After numerous attempts, a crystal soaked in 10 mM $HgCl_2$ for 1 month yielded finally a useful derivative. Seven sites belonging to seven different subunits could be located in the difference Pattersons with either the RSPS program of the CCP4 (*23*) suite, or the VSFUN programs (W. G. J. Hol, University of Groningen). The best results were obtained when the noncrystallographic axis was tilted by 5° with respect to the a^*-axis of the crystals. The heavy atom sites were refined by the program MLPHARE as implemented in the CCP4 suite, yielding an R_{Cullis} of 0.67, a phasing power of 1.86, and a combined figure of merit of 0.37 for 3970 phased reflections at 4.5 Å resolution ($R_{Cullis} = \Sigma|F_{PH,\,obs} - F_{PH,\,calc}|/\Sigma|F_{PH,\,obs} - F_P|$, where $F_{PH,\,obs}$ and $F_{PH,\,calc}$ are the observed and calculated structure factors, and F_P is the native structure factor). The single isomorphous

replacement phases were improved and extended from 5 to 3.5 Å by sevenfold averaging and solvent flattening with the program DM (*24*). Phases were extended in a total of 150 cycles. The resulting maps were readily interpretable, and most of the polypeptide could be easily fitted into the density with the program 'O' (*25*). The model was refined by using X-PLOR (*26*) with noncrystallographic constraints and restraints. The model contains 92 out of 99 residues in three different fragments. Refinement with one "group temperature factor" per residue, and an overall anisotropic temperature factor, decreased the *R* factor (for 7870 reflections) to 0.23 and the R_{free} (for 903 randomly chosen reflections) to 0.38 (*26*). The root-mean-square (rms) deviations from ideality for bond lengths and angles are 0.018 Å and 2.8°, respectively. The rms deviation between Cα atoms of different subunits in the heptamer is 0.51 Å and between all atoms 0.66 Å. The only outlying residues in the Ramachandran map belong to the flexible loop comprising residues 17 to 33, which is partially disordered in the structure, and Glu^9, which is located in a tight turn. A three-dimensional–one-dimensional profile (*27*) shows a better-than-average score for a protein of this size, with no negative values (*28*). The only region with a somewhat low score is the flexible loop. A specific-sequence marker is provided by the single mercury position per subunit in the heavy atom derivative, which is close to the side chain nitrogen of Trp^{50} (there are neither histidines nor cysteines in Mt-cpn10). Crystals of selenomethionine–Mt-cpn10 were obtained but did not result in significant peaks in difference Pattersons and difference Fouriers after data collection with multiwavelength anomalous diffraction techniques. Apparently the mobility and disorder of the single Met^{22} in the flexible loop was still too large. In our structure, we see some density that must be part of this flexible loop, but the precise way in which it is connected to the remainder of the molecule is still uncertain. Figures are drawn with MOLSCRIPT and RASTER3D (*29*).

13. S. Landry *et al.*, *Nature* **364**, 255 (1993).
14. O. Kovalenko, O. Yifrach, A. Horovitz, *Biochemistry* **33**, 14974 (1994).
15. J. Kim and R. L. Modlin, in preparation.
16. S. Chen *et al.*, *Nature* **371**, 261 (1994); J. R. Harris, R. Zahn, A. Plückthun, *J. Struct. Biol.* **115**, 68 (1995); T. Langer *et al.*, *EMBO J.* **11**, 4757 (1992).
17. S. M. van der Vies *et al.*, *Nature* **368**, 654 (1994).
18. G. Fossati *et al.*, *J. Biol. Chem.* **270**, 26159 (1995).
19. W. A. Fenton *et al.*, *Nature* **371**, 614 (1994).
20. E. S. Bochkareva and A. S. Girshovich, *J. Biol. Chem.* **267**, 25672 (1992).
21. J. S. Weissman *et al.*, *Cell* **83**, 577 (1995).
22. Evidence for hydrophobic interactions between GroEL and substrate protein has been reported by several groups including M. K. Hayer-Hartl *et al.*, *EMBO J.* **13**, 3192 (1994); Z. Lin *et al.*, *J. Biol. Chem.* **270**, 1011 (1995).
23. "The CCP4 suite: Programs for protein crystallography," *Acta Crystallogr.* **D50**, 760 (1994).
24. K. Cowtan, Joint CCP4 and ESF-EACBM, *Newsl. Prot. Crystallogr.* **31**, 34 (1994).
25. T. A. Jones *et al.*, *Acta Crystallogr.* **A47**, 110 (1991).
26. A. Brünger, X-PLOR User Manual (Yale Univ. Press, New Haven, CT, 1990); *Nature* **355**, 472 (1992).
27. R. Lüthy *et al.*, *Nature* **356**, 83 (1992).
28. S. C. Mande, V. Mehra, B. R. Bloom, W. G. J. Hol, data not shown.
29. MOLSCRIPT: P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991); RASTER3D: D. J. Bacon and W. F. Anderson, *J. Mol. Graphics* **6**, 219 (1988); E. A. Merritt *et al.*, *Acta Crystallogr.* **D50**, 869 (1994).
30. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
31. Genetic Computer Group, Incorporated, copyright 1982–1992, Madison, WI, USA.
32. A. L. Hughes, *Mol. Biol. Evol.* **10**, 1343 (1993).

33. We thank the members of the Biomolecular Structure Center for support regarding synchrotron and computing aspects of these studies, in particular S. Turley, C. Verlinde, and E. Merritt. The assistance of I. Feil in site-directed mutagenesis studies is appreciated. Access to the Stanford Synchrotron Radiation Laboratories is gratefully acknowledged. This work was supported by NIH grants AI07118 and 23545, the Howard Hughes Medical Institute, the Immunology of Mycobacteria (IMMMYC) program of the United Nations Development Programme–World Bank–World Health Organization Special Program for Research and Training in Tropical Diseases, a major equipment grant from the Murdock Charitable Trust to the Biomolecular Structure Center, and by the School of Medicine of the University of Washington, Seattle, WA. The coordinates for *M. Leprae* cpn10 have been deposited with the Brookhaven Protein Data Bank.

27 November 1995; accepted 15 December 1995

Protection Against Osmotic Stress by cGMP-Mediated Myosin Phosphorylation

Hidekazu Kuwayama, Maria Ecke, Günther Gerisch, Peter J. M. Van Haastert*

Conventional myosin functions universally as a generator of motive force in eukaryotic cells. Analysis of mutants of the microorganism *Dictyostelium discoideum* revealed that myosin also provides resistance against high external osmolarities. An osmo-induced increase of intracellular guanosine 3′,5′-monophosphate was shown to mediate phosphorylation of three threonine residues on the myosin tail, which caused a relocalization of myosin required to resist osmotic stress. This redistribution of myosin allowed cells to adopt a spherical shape and may provide physical strength to withstand extensive cell shrinkage in high osmolarities.

Cells exposed to osmotic stress can avoid dehydrative collapse either by using a cell wall (*1*) or by increasing the intracellular osmotic potential by synthesis of small molecules like glycerol, uptake of ions, or discharge of water (*2*). Actin and some of its binding proteins are required to resist high osmotic stress in yeast (*3*). The other major component of the cytoskeleton, myosin II or conventional myosin, exerts motive force by interacting with actin filaments (*4*). *Dictyostelium discoideum* is a microorganism with a small haploid genome from which mutants altered in cytoskeletal proteins have been isolated (*5–7*). Studies of mutants lacking myosin II heavy chain (mhc^-) reveal that this form of myosin is essential for cytokinesis, capping of cell surface lectin receptors, and normal cell motility and chemotaxis (*5*, *8*, *9*). Here we have used *D. discoideum* to investigate the role of myosin II in protecting amoeboid cells from high osmotic pressure.

Wild-type XP55 cells resisted an osmotic shock of 300 mM glucose for ~30 min; 50% of the cells died after a shock of about 60 min (Fig. 1A). In contrast, mhc^- cells were very sensitive to osmotic stress with a 50% reduction of cell viability after 5 to 10 min (Fig. 1A). Cells of the mhc^- mutant that had been transfected with complementary DNA (cDNA) encoding normal myosin II heavy chains (mhc^{wt}) showed a viability comparable to that of the wild-type XP55 (Fig. 1A).

Myosin II filaments bind to actin filaments more effectively than myosin monomers (*10*, *11*). Phosphorylation of myosin II at three threonine residues of its tail region inhibits filament formation (*10*, *12*) and thus weakens the interaction with actin filaments. Mutant mhc^{AAA} produces a myosin II heavy chain in which the phosphorylatable threonines are replaced by alanine residues. As a consequence of this mutation, myosin II exists predominantly in the filamentous state, and the assembly and disassembly rate are probably strongly reduced (*9*, *10*). These mhc^{AAA} cells showed the same sensitivity to high concentrations of glucose as mhc^- cells (Fig. 1A), which suggests that phosphorylation of myosin II at its tail was required to protect cells against osmotic stress (*13*).

Guanosine 3′,5′-monophosphate (cGMP) acts as a universal second messenger in eukaryotic cells (*14*). In *D. discoideum*, cGMP levels increase upon stimulation with the chemoattractant adenosine 3′,5′-monophosphate (cAMP) (*15*). Intracellular cGMP regulates the assembly and disassembly of myosin filaments by inducing the phosphorylation of the threonine residues (*16*). Osmotic stress also activates guanylyl cyclase in wild-type *D. discoideum* cells; the cGMP concentration increased after 1 min and reached a peak at ~10 min after the onset of stimulation (*17*) (Fig. 1B). A transient accumulation of cGMP levels upon addition of 300 mM glucose was also observed in wild-type XP55 and in the mhc^{wt}, mhc^-, and mhc^{AAA} strains (Fig. 1B). No increase of cGMP levels was found in the nonchemotactic mutant KI-8 in which gua-

H. Kuwayama and P. J. M. Van Haastert, Department of Biochemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, Netherlands.
M. Ecke and G. Gerisch, Max-Planck-Institut für Biochemie, D-82143 Martinsried, Germany.

*To whom correspondence should be addressed.