MAD data set was collected at the Advanced Light Source in 2 h. Diffraction data were processed with DENZO and SCALEPACK[22]. The program SOLVE[23] was used to perform the search for heavy atoms and for phasing. Solvent flattening and histogram matching with program DM[24] were used to improve the map quality. The initial model was built with program O[25] and yielded a continuous polypeptide chain between residues 5 and 262. The model was refined with package CNS[26] against the maximum-likelihood target[27] with bulk-solvent correction. Crystallographic data are summarized in Table 1.

Received 4 September; accepted 27 November 1998.

1. Doige, C. A. & Ames, G. F.-L. ATP-dependent transport systems in bacteria and humans: relevance to cystic fibrosis and multidrug resistance. *Annu. Rev. Microbiol.* **47,** 291–319 (1993).
2. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277,** 1453–1474 (1997).
3. Nikaido, K., Liu, P.-Q. & Ames, G.F.-L. Purification and characterization of HisP, the ATP-binding subunit of a traffic ATPase (ABC transporter), the histidine permease of *Salmonella typhimurium*. Solubilization, dimerization, and ATPase activity. *J. Biol. Chem.* **272,** 27745–27752 (1997).
4. Liu, C. E. & Ames, G. F.-L. Characterization of transport through the periplasmic histidine permease using proteoliposomes reconstituted by dialysis. *J. Biol. Chem.* **272,** 859–866 (1997).
5. Liu, C. E., Liu, P.-Q. & Ames, G. F.-L. Characterization of the adenosine triphosphatase activity of the periplasmic histidine permease, a traffic ATPase (ABC transporter). *J. Biol. Chem.* **272,** 21883–21891 (1997).
6. Baichwal, V., Liu, D. & Ames, G.F.-L. The ATP-binding component of a prokaryotic traffic ATPase is exposed to the periplasmic (external) surface. *Proc. Natl Acad. Sci. USA* **90,** 620–624 (1993).
7. Petronilli, V. & Ames, G.F.-L. Binding protein-independent histidine permease mutants: uncoupling of ATP hydrolysis from transmembrane signaling. *J. Biol. Chem.* **266,** 16293–16296 (1991).
8. Liu, P.-Q. & Ames, G.F.-L. *In vitro* disassembly and reassembly of an ABC transporter, the histidine permease. *Proc. Natl Acad. Sci. USA* **95,** 3495–3500 (1998).
9. Kerppola, R. E., Shyamala, V., Klebba, P. & Ames, G.F.-L. The membrane-bound proteins of periplasmic permeases form a complex: identification of the histidine permease HisQMP complex. *J. Biol. Chem.* **266,** 9857–9865 (1991).
10. Story, R. M. & Steitz, T. A. Structure of the recA protein-ADP complex. *Nature* **355,** 318–324 (1992).
11. Abrahams, J. P., Leslie, A. G., Lutter, R. & Walker, J. E. Structure at 2.8 Å resolution of F_1-ATPase from bovine heart mitochondria. *Nature* **370,** 621–628 (1994).
12. Saraste, M., Sibbald, P. R. & Wittinghofer, A. The P-loop—a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15,** 430–434 (1990).
13. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. Distantly related sequences in the α and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1,** 945–951 (1982).
14. Pai, E. F. *et al.* Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation. *Nature* **341,** 209–214 (1989).
15. Kim, S.-H., Prive, G. G. & Milburn, M. V. in *Handbook of Experimental Pharmacology. GTPases in Biology* (eds Dickey, B. F. & Birnbaumer, L.) 177–194 (Springer, Berlin, 1993).
16. Pai, E. F. *et al.* Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* **9,** 2351–2359 (1990).
17. Shyamala, V., Baichwal, V., Beall, E. & Ames, G. F.-L. Structure-function analysis of the histidine permease and comparison with cystic fibrosis mutations. *J. Biol. Chem.* **266,** 18714–18719 (1991).
18. Mimura, C. S., Admon, A., Hurt, K. A. & Ames, G. F.-L. The nucleotide-binding site of HisP, a membrane protein of the histidine permease. Identification of amino acid residues photoaffinity-labeled by 8-azido ATP. *J. Biol. Chem.* **265,** 19535–19542 (1990).
19. Ames, G. F.-L., Mimura, C., Holbrook, S. & Shyamala, V. Traffic ATPases: a superfamily of transport proteins operating from *Escherichia coli* to humans. *Adv. Enzymol.* **65,** 1–47 (1992).
20. Bianchet, M. A., Ko, Y. H., Amzel, L. M. & Pedersen, P. L. Modeling of nucleotide binding domains of ABC transporter proteins based on a F1-ATPase/recA topology: structural model of the nucleotide binding domains of the cystic fibrosis transmembrane conductance regulator (CFTR). *J. Bionerg. Biomemb.* **29,** 503–524 (1997).
21. Cheng, S. H. *et al.* Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* **63,** 827–834 (1990).
22. Otwinowski, Z. in *Data Collection and Processing* (eds Saway, L., Isaacs, N. & Bailey, S.) 56–62 (SERC Daresbury Laboratory, Warrington, UK, 1993).
23. Terwilliger, T. C., Kim, S.-H. & Eisenberg, D. Generalized method of determining heavy-atom positions using the difference Patterson function. *Acta Crystallogr. A* **43,** 1–5 (1987).
24. Cowtan, K. D. Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. *Acta Crystallogr. D* **49,** 148–157 (1993).
25. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
26. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1993).
27. Adams, P. D., Pannu, N. S., Read, R. J. & Brunger, A. T. Cross-validated maximum likelihood enhance crystallograhic simulated annealing refinement. *Proc. Natl Acad. Sci. USA* **94,** 5018–5023 (1997).
28. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thronton, J. M. Procheck—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291 (1993).
29. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24,** 946–950 (1991).
30. Merritt, E. A. & Bacon, D. J. Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277,** 505–524 (1997).

Acknowledgements. We thank J. Jancarik for crystallization of native HisP crystals, and D. S. King for carrying out mass-spectroscopy analysis on the native and Se-Met HisP proteins. The MAD data for this study were collected at the Lawrence Berkeley National Laboratory in the Macromolecular Crystallography Facility at beamline 5.0.2 in the Advanced Light Source. This facility is principally funded by the Office of Biological and Environmental Research of the US Department of Energy Offices with contributions from Lawrence Berkeley National Laboratory, Amgen, Roche Biosciences, the University of California, Berkeley, and Lawrence Livermore National Laboratory. This work has been supported by grants from the Health Effects and Life Sciences Research Division, Office of Biosciences and Environmental Research, Office of Energy Research of Department of Energy (to S.-H.K.) and from the NIH (to G.F.-L.A.).

Correspondence and requests for materials should be addressed to S.-H.K. (e-mail: shkim@lbl.gov). The atomic coordinate has been deposited at the Brookhaven Protein Databank under code 1b0u.

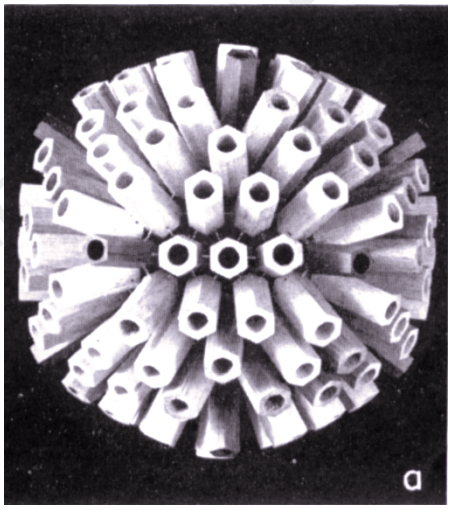
erratum

Visible viruses

Martin Kemp

Nature **396,** 123 (1998)

Readers of some copies of the 12 November issue of *Nature* will have found nothing but a black rectangle in the 'Science and Image' section where Fig. 2 was meant to be. What they should have seen is "a neat wooden model" of a herpes virus, as shown here. □



correction

Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor

Mark Paetzel, Ross E. Dalbey & Natalie C. J. Strynadka

Nature **396,** 186–190 (1998)

We omitted to give details of the coordinates of the bacterial signal peptidase : these have been deposited with the Protein Data Bank (code 1B12). Because our Letter was submitted before the 1 October 1998 deadline (see editorial in *Nature* **394**, 105; 1998), these coordinates will be held at the PDB for one year before release. □

Visible viruses

The structure of viruses was for a long time an enigma. It took an amalgam of techniques, especially the rapidly burgeoning field of electron microscopy, to reveal the quasi-symmetrical nature of viral architecture.

Martin Kemp

The exciting disclosure of the polyhedral structure of viruses during the late 1950s and early 1960s presents a perfect example of collaborative interaction between different methods of experimentation, observation, analysis and visualization, at a stage when no one method could alone deliver all the answers, because each was operating at the very margins of credibility given the immense problems of concentrating and purifying viruses. Crude X-ray diffraction, chemical analysis and shadow electron microscopy, especially with plant viruses, had begun to hint at the fascinating structures that might be involved in viral architecture, but direct viewing remained frustratingly out of reach.

The introduction of the technique of negative staining in electron microscopy by Sydney Brenner and Robert Horne in 1959 opened a visual door onto the beautiful structures and symmetries displayed in the sub-microscopic world of viruses. The method involves staining of the interstices in the structure by embedding the viruses in potassium phosphotungstate, an electron-dense material. Suitably prepared, stained and mounted, the rod-like tobacco mosaic virus and the 'spherical' turnip yellow virus stole suggestively into view, regular in configuration yet tantalizingly elusive in the precise details of their structural geometry.

The technique rapidly proved its efficacy, providing ever better definition of the symmetrical properties of these and other viruses. The improved images stimulated the building of three-dimensional models that were consistent with the emerging patterns — often still unresolved in their smaller details — in creative dialogue with the X-ray diffraction procedures that had originally disclosed the existence of symmetries. Working in Glasgow in 1960, Peter Wildy, Willie Russell and Bob Horne obtained particularly revealing electron micrographs of the negatively stained herpes simplex virus[1]. When devoid of their floppy envelopes, the protruding prisms of the morphological units or capsomeres appeared to be arranged in the 5–3–2 symmetry of an icosahedron. The capsomeres were hexameric in cross-section, with the exception of 12 at the apices that were pentameric (Fig. 1). A neat wooden model served to demonstrate the hypothetical structure (Fig. 2).

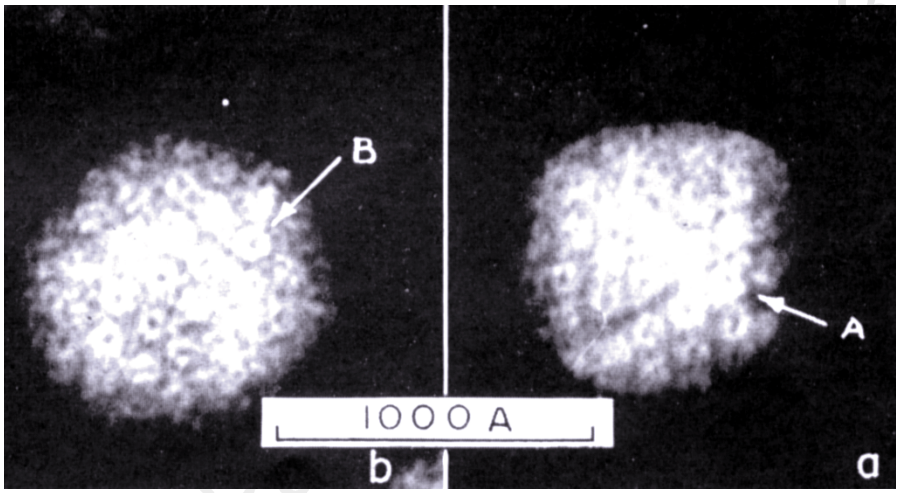


Figure 1 **Electron micrograph of herpes viruses with a five-sided unit (A) and a six-sided unit (B). (From Fig. 5 of ref. 1.)**

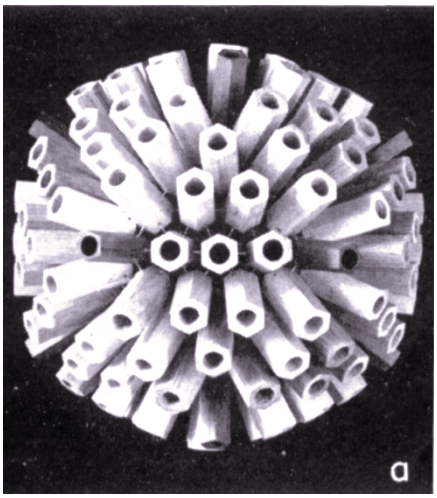


Figure 2 **Wooden model of a herpes virus with 150 hexagonal and 12 pentagonal prisms. (From Fig. 8 of ref. 1.)**

The most extended investigation of the geometrical parameters of the architecture was subsequently undertaken by Donald Caspar and Aaron Klug, as detailed at the Cold Spring Harbor Symposium in 1962. They proposed that the icosahedral regularity of the capsid and the less regular distribution of the capsomeres sometimes revealed by the electron microscope pictures could be reconciled by the concept of quasi-symmetry. In a companion article, Wildy and Douglas Watson drew parallels with the families of polyhedra characterized by Buckminster Fuller, the American visionary architect of geodesic domes whose name is now indelibly associated with C_{60} (buckminsterfullerene). As with the modelling of C_{60} in 1985, Fuller's geodesic domes proved a vital stimulus to scientific visualization. The design principles for domes, viruses and fullerenes alike combined structural economy with operational efficiency.

The story of the disclosure of the regular structure of viruses combines all the best elements of a scientific detective story. There are related bodies of visual evidence that each hover on the border of intelligibility; there is elegant speculation shaped by the search for coherent order and spiced with aesthetic intuition; and, not least, there are the empirical 'cook-book' procedures for preparation and staining to reveal the hidden secrets.

Not for nothing did Wildy and Watson quote Lewis Carroll in their contribution to the 1962 symposium:

"You boil it with sawdust: you salt it with glue:
You condense it with locusts and tape:
Still keeping one principle object in view —
To preserve its symmetrical shape."

And, as always, there is the question of 'whodunnit'. The suspects in this case remain, on the one hand, self-organization according to rules of mechanical necessity and, on the other, the organizing dictats of the nucleic acids. Or do they share joint responsibility? □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.
e-mail: martin.kemp@trinity.oxford.ac.uk

1. Wildy, P., Russell, W. C. & Horne, R. W. *Virology* **12,** 204–222 (1960).

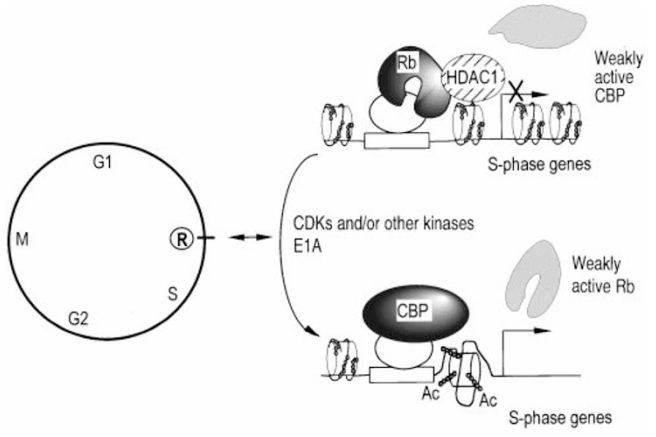


Figure 4 Hypothetical function of CBP HAT activation near the G1/S transition. At the restriction point (R), Rb is inactivated. Concomitantly, CBP is phosphorylated and activated; both actions are mimicked by E1A in transformed cells. Target genes go from a repressed state, in which the chromatin has a closed structure due to the effects of histone deacetylases recruited to the promoter by Rb, to an active state, in which the chromatin has an open configuration due to the histone acetyltransferase activity of CBP; acetylated core histone tails are designated by an Ac tag.

Methods

Plasmids. pGST–CBP: a *Bam*H1 insert from pRSV CBP (gift from R. Goodman) was inserted in-frame into pGEX-2TK. A stop codon was corrected by replacing the *Afl*II–*Bst*107 fragment with the corresponding sequence from pBSFL2CBP (gift from T. Kouzarides). Details of other constructions are available on request.

Metabolic labelling. NIH3T3 cells were stably transfected with pRSV-E1A 12s (E1A)[20] or a neomycin control vector and assayed for E1A expression (data not shown). For metabolic labelling, 10^6 NIH3T3 cells were incubated for 4 h with 0.5 mCi ml^{-1} ^{32}P inorganic phosphate (Amersham) in phosphate-free medium (Sigma) before immunoprecipitation of CBP.

Immunoprecipitation and western blotting. Stringent immunoprecipitations were performed using standard procedures; cells were lysed in RIPA buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1% N-P40, 0.5% sodium deoxycholate, 0.1% SDS, 1 mM EDTA). SRC-1 and P/CAF were undetectable in the immunoprecipitates (data not shown). Mild immunoprecipitation was performed as described[10]. All buffers contained protease (Boehringer) and phosphatase (Sigma) inhibitors. Western blotting was done using standard procedures and visualized using an ECL$^+$ kit (Amersham). Anti-CBP antibodies used were the A22 antibody (Santa Cruz) for immunoprecipitation and the NM11 antibody (Pharmingen) for western blotting.

***In vitro* phosphorylation and phosphatase treatment.** Cyclin E/Cdk2 protein kinase, expressed in baculovirus-infected Sf9 cells, was purified by standard chromatographic procedures. GST proteins were purified as described[29] and dialysed against TBS-G (20 mM Tris, pH 8.0, 150 mM NaCl, 10% glycerol). GST–CBP contained full-length CBP (as assessed by western blot analysis with anti-C-terminus and anti-N-terminus antibodies; data not shown). His-Rb was prepared using the Qiagen protocol. Recombinant proteins were *in vitro* phosphorylated by incubation with 50 ng cyclin E–Cdk2, 100 μM ATP and [γ-^{32}P]ATP (100 mCi $mmol^{-1}$ final specific activity) for 45 min at 30 °C in 30 μl buffer containing 25 mM Tris, pH 7.5, 0.1 mM $NaVO_4$, 0.1 mM EGTA, 10 mM magnesium acetate, 0.04 mM DTT, 0.1 mM $ZnSO_4$ and protease inhibitors. Phosphatase treatment was performed using 400 U of lambda protein phosphatase (Biolabs) for 30 min at 30 °C.

HAT assay. HAT assays[30] were done using a synthetic peptide (Chiron) corresponding to the first 24 amino acids of histone H4 coupled through a linker sequence to a biotin molecule.

Received 24 July; accepted 2 September 1998.

1. Wang, H. G., Moran, E. & Yaciuk, P. E1A promotes association between p300 and pRB in multimeric complexes required for normal biological activity. *J. Virol.* **69,** 7917–7924 (1995).
2. Dyson, N. & Harlow, E. Adenovirus E1A targets key regulators of cell proliferation. *Cancer Surv.* **12,** 161–195 (1992).
3. Wang, H. G. *et al.* Identification of specific adenovirus E1A N-terminal residues critical to the binding of cellular proteins and to the control of cell growth. *J. Virol.* **67,** 476–488 (1993).
4. Kouzarides, T. Transcriptional control by the retinoblastoma protein. *Semin. Cancer Biol.* **6,** 91–98 (1998).
5. Weintraub, S. J., Prater, C. A. & Dean, D. C. Retinoblastoma protein switches the E2F site from positive to negative element. *Nature* **358,** 259–261 (1992).
6. Chrivia, J. C. *et al.* Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **265,** 855–859 (1993).
7. Lundblad, J. R., Kwok, R. P., Laurance, M. E., Harter, M. L. & Goodman, R. H. Adenoviral E1A-associated protein p300 as a functional homologue of the transcriptional co-activator CBP. *Nature* **374,** 85–88 (1995).
8. Kwok, R. P. *et al.* Nuclear protein CBP is a co-activator for the transcription factor CREB. *Nature* **370,** 223–226 (1994).
9. Arany, Z., Newsome, D., Oldread, E., Livingston, D. M. & Eckner, R. A family of transcriptional adaptor proteins targeted by the E1A oncoprotein. *Nature* **374,** 81–84 (1995).
10. Magnaghi-Jaulin, L. *et al.* Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature* **391,** 601–605 (1998).
11. Brehm, A. *et al.* Retinoblastoma protein recruits histone deacetylases to repress transcription. *Nature* **391,** 597–601 (1998).
12. Luo, R. X., Postigo, A. A. & Dean, D. C. Rb interacts with histone deacetylase to repress transcription. *Cell* **92,** 463–473 (1998).
13. Bannister, A. J. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384,** 641–643 (1996).
14. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87,** 953–959 (1996).
15. Weinberg, R. A. The retinoblastoma gene and gene product. *Cancer Surv.* **12,** 43–57 (1992).
16. Kitabayashi, I. *et al.* Phosphorylation of the adenovirus E1A-associated 300 kDa protein in response to retinoic acid and E1A during the differentiation of F9 cells. *EMBO J.* **14,** 3496–3509 (1995).
17. Yaciuk, P. & Moran, E. Analysis with specific polyclonal antiserum indicates that the E1A-associated 300-kDa product is a stable nuclear phosphoprotein that undergoes cell cycle phase-specific modification. *Mol. Cell. Biol.* **11,** 5389–5397 (1991).
18. Perkins, N. D. *et al.* Regulation of NF-κB by cyclin-dependent kinases associated with the p300 coactivator. *Science* **275,** 523–527 (1997).
19. Borrow, J. *et al.* The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nature Genet.* **14,** 33–41 (1996).
20. Trouche, D. & Kouzarides, T. E2F1 and E1A(12S) have a homologous activation domain regulated by RB and CBP. *Proc. Natl Acad. Sci. USA* **93,** 1439–1442 (1996).
21. Nevins, J. R., DeGregori, J., Jakoi, L. & Leone, G. Functional analysis of E2F transcription factor. *Methods Enzymol.* **283,** 205–219 (1997).
22. Goldman, P. S., Tran, V. K. & Goodman, R. H. The multifunctional role of the co-activator CBP in transcriptional regulation. *Recent Prog. Horm. Res.* **52,** 103–119 (1997).
23. Eckner, R., Yao, T. P., Oldread, E. & Livingston, D. M. Interaction and functional collaboration of p300/CBP and bHLH proteins in muscle and B-cell differentiation. *Genes Dev.* **10,** 2478–2490 (1996).
24. Puri, P. L. *et al.* Differential roles of p300 and PCAF acetyltransferases in muscle differentiation. *Mol. Cell* **1,** 35–45 (1997).
25. Korzus, E. *et al.* Transcription factor-specific requirements for coactivators and their acetyltransferase functions. *Science* **279,** 703–707 (1998).
26. Ramirez, S., Ait-Si-Ali, S., Robin, P., Trouche, D. & Harel-Bellan, A. The CREB-binding protein (CBP) cooperates with the serum response factor for transactivation of the c-fos serum response element. *J. Biol. Chem.* **272,** 31016–31021 (1997).
27. Swope, D. L., Mueller, C. L. & Chrivia, J. C. CREB-binding protein activates transcription through multiple domains. *J. Biol. Chem.* **271,** 28138–28145 (1996).
28. Martinez-Balbas, M. A. *et al.* The acetyl-transferase activity of CBP stimulates transcription. *EMBO J.* **17,** 2886–2893 (1998).
29. Groisman, R. *et al.* Physical interaction between the mitogen-responsive serum response factor and myogenic bHLH proteins. *J. Biol. Chem.* **271,** 5258–5264 (1996).
30. Ait-Si-Ali, S., Ramirez, S., Robin, P., Trouche, D. & Harel-Bellan, A. A rapid and sensitive assay for histone acetyl-transferase activity. *Nucleic Acids Res.* **26,** 3869–3870 (1998).

Acknowledgements. We thank Z. Mishal and A. Vervisch for help with cell-cycle analysis; T. Kouzarides, L. Meijer and D. A. Lawrence for the gift of materials; F. Dautry for critical reading of the manuscript; and A. Damany for her support. This work was supported by grants from the Ligue Nationale contre le Cancer, the Comité des Yvelines, the Comité de l'Essone and the Comité du Val de Marne, from the Association pour la Recherche sur le Cancer and from the Groupement des Entreprises Françaises dans la Lutte contre le Cancer. S.A.-S.-A. was awarded a fellowship from the Comité de la Haute-Saône; S.R., a travel award from the Colombian Government (Colciencias); F.-X.B., a fellowship from the Agence Nationale pour la Recherche sur le Sida; L.M.-J. and F.D., fellowships from the Comité de l'Essone.

Correspndence and requests for materials should be addressed to A.H.-B. (e-mail: ahbellan@vjf.cnrs.fr).

Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor

Mark Paetzel*, Ross E. Dalbey† & Natalie C. J. Strynadka*

** Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver V6T 1Z3, British Columbia, Canada*
† Department of Chemistry, The Ohio State University, Columbus, Ohio 43210, USA

The signal peptidase (SPase) from *Escherichia coli* is a membrane-bound endopeptidase with two amino-terminal transmembrane segments and a carboxy-terminal catalytic region which resides in the periplasmic space[1]. SPase functions to release proteins that have been translocated into the inner membrane from the cell

interior, by cleaving off their signal peptides[1]. We report here the X-ray crystal structure of a catalytically active soluble fragment of *E. coli* SPase (SPase Δ2–75)[2,3]. We have determined this structure at 1.9 Å resolution in a complex with an inhibitor, a β-lactam (5*S*,6*S* penem)[4,5], which is covalently bound as an acyl-enzyme intermediate to the γ-oxygen of a serine residue at position 90, demonstrating that this residue acts as the nucleophile in the hydrolytic mechanism of signal-peptide cleavage. The structure is consistent with the use by SPase of Lys 145 as a general base in the activation of the nucleophilic Ser 90, explains the specificity requirement at the signal-peptide cleavage site, and reveals a large exposed hydrophobic surface which could be a site for an intimate association with the membrane. As enzymes that are essential for cell viability, bacterial SPases present a feasible antibacterial target[4–6]: our determination of the SPase structure therefore provides a template for the rational design of antibiotic compounds.

The *E. coli* SPase Δ2–75 structure has a mainly β-sheet protein fold, consisting of two large antiparallel β-sheet domains (termed I and II and coloured green and blue, respectively, in Fig. 1), two small 3_{10}-helices (consisting of residues 246–250 and 315–319), and one small α-helix (residues 280–285). There is one disulphide bond, as was found in earlier biochemical studies[7], between Cys 170 and Cys 176. This bond is located immediately before a β-turn in the domain II β-sheet (Fig. 1). In addition, an extended β-ribbon (residues 107–122, coloured purple in Fig. 1) protrudes from domain I, together with the N-terminal strand, giving the SPase Δ2–75 molecule an overall conical shape with rough dimensions of 60 × 40 × 70 Å (Figs 1, 2).

Sequence alignments indicate that highly conserved regions of primary sequence within the prokaryotic and eukaryotic SPases[1] reside within domain I of the *E. coli* SPase structure, whereas the two insertions representing the extended β-ribbon (residues 107–122) and domain II are variably present from species to species. In addition, domain I shares structural similarities with UmuD′ protease[8], the proteolytic domain of a self-cleaving repressor protein involved in the 'SOS' DNA-repair response in *E. coli*. Although the overall mainchain connectivity in UmuD′ and domain I of SPase differs in some regions, 68 common Cα atoms can be superimposed with a root mean square (r.m.s.) deviation of 1.6 Å. Domain II and the extended β-ribbon of SPase have no structural counterparts in UmuD′. Domain I, containing all of the essential and conserved catalytic elements, represents a new protease structural motif that is likely to be conserved from bacteria to man.

A large, unusually exposed hydrophobic surface extends across the SPase Δ2–75 molecule and includes the substrate-binding site and catalytic centre (labelled S1, S3 and Ser 90 in Fig. 2). The residues contributing to the hydrophobic character of this surface

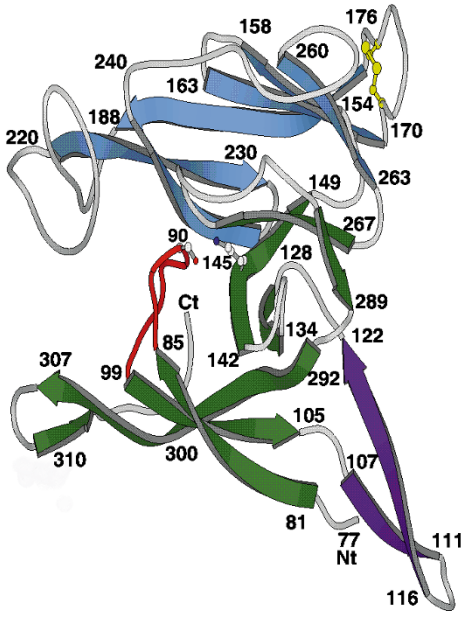


Figure 1 A ribbon diagram[28] showing the general fold of SPase Δ2–75. The domain I β-sheet (the conserved catalytic core) is shown in green and the domain II β-sheet in blue. The β-hairpin extension protruding from domain I is shown in purple. The active-site residues Ser 90 and Lys 145 are labelled 90 and 145. The loop containing the nucleophilic Ser 90 is in red. The disulphide bond between Cys 170 and Cys 176 is shown in yellow. The inhibitor is not shown for clarity. Likewise, some small β-strands and helices are shown as random coils for clarity. The antiparallel β-strands consisting of residues 81–85, 99–105, 292–307 and 312–314 form a large exposed hydrophobic surface (the proposed membrane-association surface).

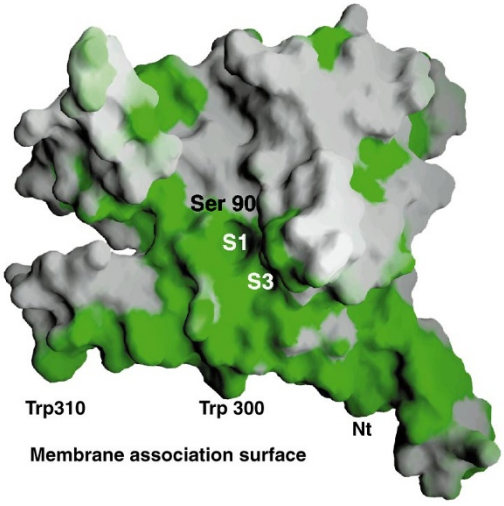


Figure 2 A GRASP[29] representation of the molecular surface of SPase Δ2–75. The view is the same as in Fig. 1. Green represents exposed hydrophobic surfaces. The substrate-binding sites S1 and S3 are labelled, as is Ser 90 of the active site. Trp 300, Trp 310, and the N terminus (Nt) are labelled along the large hydrophobic surface, the proposed membrane-association surface.

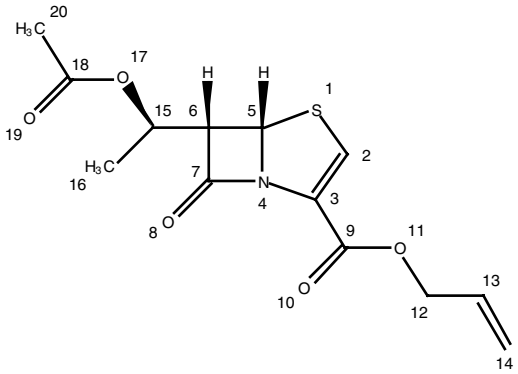


Figure 3 Structure of the β-lactam-type inhibitor allyl (5*S*,6*S*)-6-[(*R*)-acetoxyethyl]-penem-3-carboxylate[4,5].

protrude from the main β-sheet of domain I and include Phe 79, Ile 80, Tyr 81, Phe 100, Leu 102, Trp 300, Met 301, Phe 303, Trp 310, Leu 314, Leu 316, and Ile 319. On the basis of our observations of the structure we suggest that, *in vivo*, the membrane-anchored N-terminal strand and the associated β-ribbon, from residues 106 to 124 (Fig. 1), would bend in the appropriate manner to allow the exposed hydrophobic surface of SPase to insert into the membrane lipid bilayer, presumably optimizing contact with the signal-peptide cleavage site. This proposal is consistent with earlier experiments that showed that the detergent Triton-X100 is essential for optimal activity of SPase Δ2–75 (ref. 3), as well as for optimal growth of the SPase Δ2–75 crystals[9]. Recent biophysical studies[10] have revealed that SPase Δ2-75 inserts into the outer leaflet of the *E. coli* inner membrane. In addition, it has been suggested[11] that Trp 300 and Trp 310 (Figs 1, 2) are essential for the catalytic activity of *E. coli* SPase. This result is intriguing given the distance (>20 Å) between these residues and the active site and their location on the hydrophobic surface, the proposed membrane-association surface (Figs 1, 2). Tryptophans and other aromatic residues are commonly found at membrane–protein interfaces[12].

Although bacterial SPases are not inhibited by standard protease inhibitors, they are inhibited by β-lactam compounds with 5*S* stereochemistry[4,5]. We have determined the structure of SPase Δ2–75 in the presence of a 5*S*,6*S* β-lactam (penem), an SPase inhibitor (Fig. 3)[4,5]. The electron density shows a covalent bond between SPase Ser 90 Oγ and the carbonyl carbon (C7) of the inhibitor, with the four-membered β-lactam ring being cleaved between C7 and N4 (Fig. 4). This is, to our knowledge, the first direct evidence for the role of Ser 90 Oγ as the acylating nucleophile in catalysis. The structure shows that the Ser 90 Oγ attacks the *si*-face of the β-lactam amide bond, a peptide-bond analogue. This indicates that SPase may be unique among serine-dependent hydrolases, including the serine proteases[4,13] and the group 2b β-lactamases[14], which prefer a *re*-face attack. A *si*-face nucleophilic attack by *E. coli* SPase was predicted previously on the basis of stereochemical requirements of several inhibitory compounds[4].

The main-chain amide of Ser 90 forms a strong hydrogen bond (of length 2.9 Å) with the carbonyl oxygen (O8) of the cleaved β-lactam ring (Fig. 4). This indicates that the Ser 90 amide might contribute to the formation of an 'oxyanion hole', lending electrophilic assistance by stabilizing the tetrahedral transition-state intermediate. There appears to be no suitably positioned second main-chain or sidechain amide that could contribute to the oxyanion hole (as is found in the group 2b β-lactamases[14] and the serine proteinases[15]). However, the Ser 88 side chain could potentially participate in such an interaction by a simple rotation from the observed χ_1 of −54° (Fig. 4) to a value of +60° (Fig. 5). This interaction is prevented in the inhibitor complex by an unfavourable van der Waals contact between the Ser 88 Oγ in the +60° conformation and the S1 and C15 atoms of the inhibitor (Fig. 4). The Ser 88 side chain has the highest temperature factors in the active-site region, indicating that it is not in an optimal environment in the inhibitor complex. The contribution of a serine hydroxyl to an oxyanion hole has been seen previously in lipolytic enzymes such as cutinase[16].

The Lys 145 Nζ position is fixed relative to Ser 90 Oγ by hydrogen bonds to Ser 278 Oγ (bond length 2.9 Å) and to the carbonyl oxygen (O10) of the inhibitor side chain (bond length 2.9 Å) (Fig. 4). The Nζ of Lys 145 is 2.9 Å away from the Ser 90 Oγ and is the only titratable group in the vicinity of the active-site nucleophile (Fig. 4). The next closest ionizable group, 7.5 Å away from Ser 90 Oγ, is Asp 280 which is held in place by a strong salt bridge to Arg 282 (Fig. 4). Thus, the ϵ-amino group of Lys 145 is suitably positioned to act as the general base in both acylation and deacylation steps of catalysis. It appears as though the inhibitor has displaced the deacylating water, as no water molecules are found within 5.5 Å of the covalent inhibitor link. As the co-crystals containing enzyme and inhibitor were grown weeks before the data collection, the acyl-enzyme must be extremely stable, supporting the idea that a deacylating water molecule is displaced. The side chain of Lys 145 is completely buried in this inhibitor complex (Fig. 4), in which it makes van der Waals contacts with the sidechain atoms of Tyr 143, Phe 133 and Met 270, and with the main-chain atoms of Met 270, Met 271, Gly 272 and Ala 279, all of which come from domain I. The

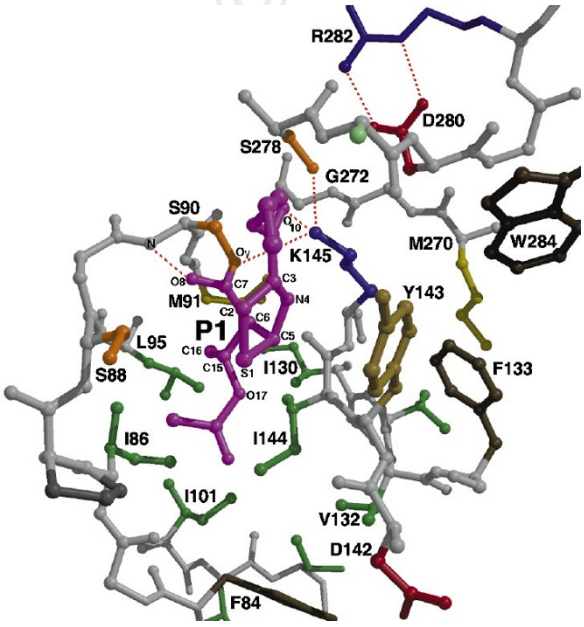


Figure 4 A ball-and-stick representation[30] of the active-site residues of SPase Δ2–75. The β-lactam (5*S*,6*S* penem) inhibitor[4,5] shown in Fig. 3 and in this figure (purple) is covalently bound to the Oγ of Ser 90, with the carbonyl oxygen (O8) of the cleaved β-lactam (the bond between C7 and N4 has been cleaved) sitting in the oxyanion hole formed by the main-chain nitrogen of Ser 90 (S90). The methyl group (C16) of the inhibitor, labelled P1, sits in the S1 substrate-binding site.

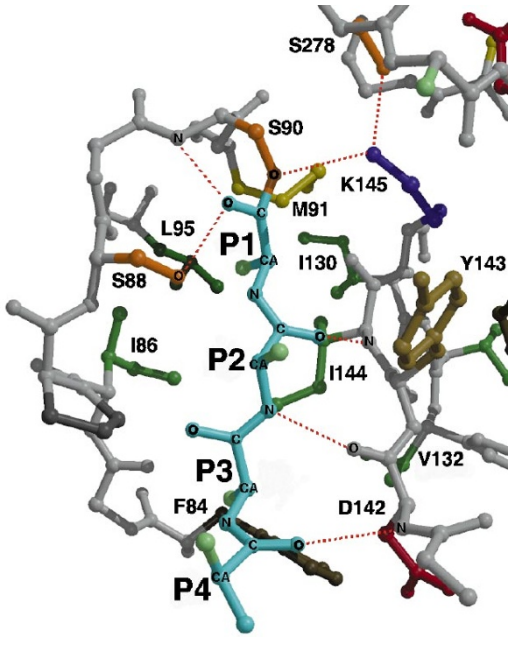


Figure 5 A ball-and-stick representation[30] of the active-site residues of SPase Δ2–75 with the P1–P4 residues of an acylated peptide substrate (Ala-Ala-Ala-Ala) modelled into the bindings sites S1–S4. The observed positions of the methyl group (C16) and the carbonyl oxygen (O8) of the inhibitor (Fig. 4) were used as a guide.

hydrophobic environment surrounding the Lys 145 ϵ-amino group is probably essential for lowering its pK_a so that it can stay in the deprotonated state required for its function as the general base[17,18].

A typical *E. coli* signal peptide consists of a positively charged N terminus, an inner hydrophobic core, and a C-terminal cleavage-recognition sequence containing small uncharged residues at the P1 (−1) and P3 (−3) sites[19,20]. Alanine residues are the most common residues at the −1 and −3 positions, giving rise to the so-called −1, −3 or Ala-X-Ala, rule[19,20]. The sidechain methyl group (C16) of the penem is located in the SPase substrate-binding pocket (S1) (Figs 2, 4). This methyl group is essential for the effectiveness of the inhibitor[4,5] and probably mimics the P1 (−1) (Ala) side chain of the substrate. The residues making direct van der Waals contacts with the P1 methyl group in the S1 specificity pocket are Met 91, Ile 144, Leu 95 and Ile 86 (Fig. 4).

Using the position of the inhibitor methyl group (C16) in the S1 site and of the inhibitor carbonyl group (C7, O8) in the oxyanion hole as a guide, we have modelled a tetrapeptide (poly-Ala) into the active site of SPase (Fig. 5). We needed an extended, β-strand conformation of the peptide substrate to provide both a favourable fit and β-sheet-type hydrogen bonds with the conserved β-strand containing Lys 145, supporting earlier studies which indicated that the C-terminal five to six residues of the signal peptide would adopt a β-sheet conformation[20]. This model helps to explain the cleavage-site specificity of SPase. The side chain of the P1 Ala occupies the same site as the inhibitor methyl group and the side chain of the P3 Ala points into a shallow hydrophobic depression formed by Phe 84, Ile 86, Ile 101, Val 132, Ile 144 and the Cβ of Asp 142 (the proposed substrate-specificity site S3; Figs 2, 5). Although alanine is the most common residue at the P3 site of signal peptides, larger aliphatic residues such as Val, Leu and Ile can also occur at this position. Our structure shows that the hydrophobic depression for the S3 site is broader than that for the S1 site (Figs 2, 5). The S3 site could therefore accommodate these larger residues at the P3 site of the signal-peptide substrate. The side chains of the residues at P2 and P4 point out of the active site towards the solvent (Fig. 5), consistent with the observed signal-peptide sequence variability at these positions[19].

Future modelling studies aimed at an understanding of the structure and function of the eukaryotic SPases will proceed on the basis of the conservation of primary sequence[1] within the *E. coli* SPase domain I, the catalytic core of type 1 SPases. Important issues, such as the reasons behind unique substrate specificity of mitochondrial SPases[1] and the substitution of the catalytic lysine by the more typical histidine in the endoplasmic reticulum SPases[1], can now be addressed from this first structure of an SPase. □

Methods

Data collection. The SPase Δ2–75 protein (relative molecular mass (M_r) 27,952; 249 amino-acid residues) was expressed and purified as described[9]. The crystals were grown in the presence of the inhibitor allyl (5*S*,6*S*)-6-[(*R*)-acetoxyethyl]penem-3-carboxylate[4,5] and the detergent Triton-X100. Because of the complicated methodology involved, the procedure for the crystallization of the orthorhombic crystal form of SPase Δ2–75 will be published elsewhere. The crystals belong to the orthorhombic space group $P2_12_12$ with unit-cell dimensions of $a = 110.7$ Å, $b = 113.2$ Å, $c = 99.2$ Å. The specific volume (V_m)[21] of the crystals was 2.78 Å^3 Da^{-1} for four molecules in the asymmetric unit. The fraction of crystal volume occupied by solvent was ~56%. The ethylmercury phosphate and methylmercury acetate soaks were done at concentrations of 4.9 mM and 6.1 mM for 6 and 12 h, respectively. The diffraction intensities were measured at 100 K on beamline X12C at the Brookhaven National Synchrotron Light Source (NSLS). The data were processed with the program DENZO[22].

Phase determination and refinement. We determined the crystal structure of SPase Δ2–75 by multiple isomorphous replacement with anomalous signal (MIRAS) using the phases calculated from two heavy-atom derivatives (ethylmercury phosphate and methylmercury acetate)[23]. The heavy-atom parameters were refined and phases calculated using the program MLPHARE[23]. The resulting electron-density map was greatly improved by solvent flattening, histogram matching, and non-crystallographic symmetry averaging (four molecules in the asymmetric unit) using the program DM[23]. Molecular-model building into the electron-density map was done with the program O (ref. 24) and the structure was refined using the programs XPLOR[25] and TNT[26]. Phasing and refinement statistical parameters are shown in Table 1. The most disordered regions in each of the molecules in the asymmetric unit are extended loops or hairpins near the solvent surface (residues 108–124, 170–176, 198–202 and 304–313) and are still under refinement. The N-terminal Met and Val 76 are not observed in any of the four molecules of the asymmetric unit. An error in the reported amino-acid sequence[27] was observed from the electron density and confirmed by DNA sequencing: Ala (GCT) 182 is Val (GTC) 182.

Table 1 Crystallographic data

Data collection

Data set	d_{max}* (Å)	Reflections			⟨$I/\sigma(I)$⟩	R_{merge}† (%)
		Total observed	Unique	Percent of possible		
Native	1.9	391,951	88,159	97.2	21.9	5.6
Ethylmercury phosphate	2.9	98,274	28,379	99.6	7.9	6.9
Methylmercury acetate	2.9	91,557	21,851	76.5	6.8	10.3

Phasing statistics

Derivative	Resolution (Å)	Sites	PhP‡ Acentric/centric	R_{cullis}§ Acentric/centric
Ethylmercury phosphate‖	20.0–2.9	7	1.30/1.02	0.78/0.72
Methylmercury acetate	10.0–4.0	2	0.95/0.69	0.86/0.84

Current refinement statistics

Completeness of model			R¶ (%)	R_{free}# (%)	r.m.s. deviation		B_{ave} (Å^2)
Residues	Atoms	Water molecules			Bonds (Å)	Angles (°)	
988	7,899	253	22.5	27.7	0.016	1.9	29.2

* d_{max} is the maximum resolution of measured X-ray intensities.
† $R_{merge} = \Sigma|I_{o,i} - I_{ave,i}|/\Sigma I_{ave,i}|$, where $I_{ave,i}$ is the average structure-factor amplitude of reflection I and $I_{o,i}$ represents the individual measurements of reflection I and its symmetry equivalent reflection.
‡ PhP is the phasing power $= \sqrt{\Sigma F_H^2}/\sqrt{\Sigma(|F_{PHobs}| - |F_{PHcalc}|)^2}$, where F_{PH} and F_H are the derivative and calculated heavy-atom structure factors, respectively.
§ $R_{cullis} = \Sigma||F_{PH} \pm F_P| - F_{Hcalc}|/\Sigma|F_{PH} \pm F_P|$, where F_{PH}, F_P and F_H are the derivative, native and calculated heavy-atom structure factors, respectively.
‖ An anomalous signal to 4 Å resolution was used. The overall figure of merit for both derivatives, including the anomalous signal, was 0.37–2.9 Å.
¶ $R = \Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$ (on all data 1.9–20.0 Å).
# $R_{free} = \Sigma_{hkl\subset T}(|F_o| - |F_c|)^2/\Sigma_{hkl\subset T}|F_o|^2$, where $\Sigma_{hkl\subset T}$ are reflections belonging to a test set of 10% of the data.
r.m.s., root mean square.

Received 7 July; accepted 21 September 1998.

1. Dalbey, R. E., Lively, M. O., Bron, S. & van Dijl, J. M. The chemistry and enzymology of the type 1 signal peptidases. *Protein Sci.* **6,** 1129–1138 (1997).
2. Kuo, D. W. *et al*. *Escherichia coli* leader peptidase: production of an active form lacking a requirement for detergent and development of peptide substrates. *Arch. Biochem. Biophys.* **303,** 274–280 (1993).
3. Tschantz, W. R. *et al*. Characterization of a soluble, catalytically active form of *Escherichia coli* leader peptidase: requirement of detergent or phospholipid for optimal activity. *Biochemistry* **34,** 3935–3941 (1995).
4. Allsop, A. E. *et al*. in *Anti-Infectives, Recent Advances in Chemistry and Structure-Activity Relationships* (eds Bently, P. H. & O'Hanlon, P. J.) 61–72 (R. Soc. Chem., Cambridge, 1997).
5. Black, M. T. & Bruton, G. Inhibitors of bacterial signal peptidases. *Curr. Pharm. Des.* **4,** 133–154 (1998).
6. Date, T. Demonstration by a novel genetic technique that leader peptidase is an essential enzyme in *Escherichia coli*. *J. Bacteriol.* **154,** 76–83 (1983).
7. Whitely, P. & von Heijne, G. The DsbA-DsbB system affects the formation of disulfide bonds in periplasmic but not in intramembraneous protein domains. *FEBS Lett.* **332,** 49–51 (1993).
8. Peat, T. S. *et al*. Structure of the UmuD′ protein and its regulation in response to DNA damage. *Nature* **380,** 727–730 (1996).
9. Paetzel, M. *et al*. Crystallization of a soluble, catalytically active form of *Escherichia coli* leader peptidase. *Proteins Struct. Funct. Genet.* **23,** 122–125 (1995).
10. van Klompenburg, W. *et al*. Phosphatidylethanolamine mediated insertion of the catalytic domain of leader peptidase in membranes. *FEBS Lett.* **431,** 75–79 (1998).
11. Kim, Y. T., Muramatsu, T. & Takahashi, K. Identification of Trp 300 as an important residue for *Escherichia coli* leader peptidase activity. *Eur. J. Biochem.* **234,** 358–362 (1995).
12. Landolt-Marticorena, C., Williams, K. A., Deber, C. M. & Reithmeirer, R. A. Non-random distribution of amino acids in the transmembrane segments of human type I single span membrane proteins. *J. Mol. Biol.* **229,** 602–608 (1993).
13. James, M. N. G. in *Proteolysis and Protein Turnover* (eds Bond, J. S. & Barrett, A. J.) 1–8 (Portland, Brookfield, VT, 1994).
14. Strynadka, N. C. J. *et al*. Molecular structure of the acyl-enzyme intermediate in β-lactamase at 1.7 Å resolution. *Nature* **359,** 393–400 (1992).
15. Manard, R. & Storer, A. C. Oxyanion hole interactions in serine and cysteine proteases. *Biol. Chem. Hoppe-Seyler* **373,** 393–400 (1992).
16. Nicolas, A. *et al*. Contribution of cutinase Ser 42 side chain to the stabilization of the oxyanion transition state. *Biochemistry* **35,** 398–410 (1996).
17. Paetzel, M. *et al*. Use of site-directed chemical modification to study an essential lysine in *Escherichia coli* leader peptidase. *J. Biol. Chem.* **272,** 9994–10003 (1997).
18. Paetzel, M. & Dalbey, R. E. Catalytic hydroxyl/amine dyads with serine proteases. *Trends Biochem. Sci.* **22,** 28–31 (1997).
19. von Heijne, G. Signal sequences. The limits of variation. *J. Mol. Biol.* **184,** 99–105 (1985).
20. Izard, J. W. & Kendall, D. A. Signal peptides: exquisitely designed transport promoters. *Mol. Microbiol.* **13,** 765–773 (1994).
21. Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **33,** 491–497 (1968).
22. Otwinowski, Z. in *DENZO* (eds Sawyer, L., Isaacs, N. & Baily, S.) 56–62 (SERC Daresbury Laboratory, Warrington, UK, 1993).
23. Collaborative Computational Project No. 4 The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
24. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kieldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
25. Brunger, A. T. *X-PLOR: A System for X-ray Crystallography and NMR (Version 3.1)* (Yale Univ. Press, New Haven, 1987).
26. Tronrud, D. E. Conjugate-direction minimization: an improved method for the refinement of macromolecules. *Acta Crystallogr. A* **48,** 912–916 (1992).
27. Wolfe, P. B., Wickner, W. & Goodman, J. M. Sequence of the leader peptidase gene of *Escherichia coli* and the orientation of leader peptidase in the bacterial envelope. *J. Biol. Chem.* **258,** 12073–12080 (1983).
28. Kraulis, P. G. Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24,** 946–950 (1991).
29. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and the thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11,** 281–296 (1991).
30. Meritt, E. A. & Bacon, D. J. Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277,** 505–524 (1997).

Acknowledgements. We thank SmithKlineBeecham Pharmaceuticals for penem inhibitor; R. M. Sweet for use of beamline X12C (NSLS, Brookhaven National Laboratory); G. Petsko for the ethylmercury phosphate; M. N. G. James for access to equipment for characterization of earlier crystal forms of SPase; and S. Mosimann and S. Ness for discussions. This work was supported by the Medical Research Council of Canada, the Canadian Bacterial Diseases Network of Excellence, and British Columbia Medical Research Foundation grants to N.C.J.S. M.P. is funded by an MRC of Canada post-doctoral fellowship, N.C.J.S. by an MRC of Canada scholarship, and R.E.D. by the NIH and the American Heart Association.

Correspondence and requests for materials should be addressed to N.C.J.S. (e-mail: natalie@byron.biochem.ubc.ca).

errata

Reconciling the spectrum of Sagittarius A* with a two-temperature plasma model

Rohan Mahadevan

Nature **394,** 651–653 (1998)

A misleading typographical error was introduced into the second sentence of the bold introductory paragraph of this Letter: the word "infrared" should be "inferred". □

Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence

S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M.-A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead & B. G. Barrell

Nature **393,** 537–544 (1998)

As a result of an error during film output, Table 1 was published with some symbols missing. The correct version can be found at http://www.sanger.ac.uk and is reproduced again here (following pages).

Also, in Fig. 2, we incorrectly labelled Rv0649 as *fadD37* instead of *fabD2*. Two of the genes for mycolyl transferases were inverted: Rv0129c encodes antigen 85C and not 85C′ as stated, whereas Rv3803c codes for the secreted protein MPT51 and not antigen 85C (*Infect. Immun.* **59,** 372–382; 1991); Rv3803c is now designated *fbpD*. We thank Morten Harboe and Harald Wiker for drawing this to our attention.

The sequence of Rv0746 from *M. bovis* BCG-Pasteur presented in Fig. 5b was incorrect and should have shown a 16-codon deletion instead of 29, as indicated here:

```
H37Rv.....GSGAPGGAGGAAGLWGTGGAGGAGGSSAGGGGAGGAGGAGGWLLGDGGAGGIGGAST...
..........:::::::::::::::::::::::          ::::::::::::::::::::
BCG.......GSGAPGGAGGAAGLWGTGGA----------------GGAGGWLLGDGGAGGIGGAST...
```

□