Reversible Compartmentalization of de Novo Purine Biosynthetic Complexes in Living Cells

Article *in* Science · May 2008

DOI: 10.1126/science.1152241 · Source: PubMed

4 authors, including:

Songon An
University of Maryland, Baltimore County
35 PUBLICATIONS **833** CITATIONS
SEE PROFILE

Ravindra Kumar
Merck & Co.
11 PUBLICATIONS **427** CITATIONS
SEE PROFILE

Erin D Sheets
University of Minnesota Duluth
50 PUBLICATIONS **2,889** CITATIONS
SEE PROFILE

Reversible Compartmentalization of de Novo Purine Biosynthetic Complexes in Living Cells

Songon An, *et al.*

Science **320**, 103 (2008);
DOI: 10.1126/science.1152241

In addition to the finding that bacteria subsisting on natural and synthetic antibiotics are widely distributed in the environment, these results highlight an unrecognized reservoir of multiple antibiotic-resistance machinery. Bacteria subsisting on antibiotics are phylogenetically diverse and include many organisms closely related to clinically relevant pathogens. It is thus possible that pathogens could obtain antibiotic-resistance genes from environmentally distributed super-resistant microbes subsisting on antibiotics.

References and Notes

1. C. S. Riesenfeld, R. M. Goodman, J. Handelsman, *Environ. Microbiol.* **6**, 981 (2004).
2. V. M. D'Costa, K. M. McGrann, D. W. Hughes, G. D. Wright, *Science* **311**, 374 (2006).
3. C. Walsh, *Nature* **406**, 775 (2000).
4. M. N. Alekshun, S. B. Levy, *Cell* **128**, 1037 (2007).
5. J. Davies, *Science* **264**, 375 (1994).
6. J. K. Fredrickson, H. M. Kostandarithes, S. W. Li, A. E. Plymale, M. J. Daly, *Appl. Environ. Microbiol.* **66**, 2006 (2000).
7. K. A. McAllister, H. Lee, J. T. Trevors, *Biodegradation* **7**, 1 (1996).
8. Y. Kameda, E. Toyoura, Y. Kimura, T. Omori, *Nature* **191**, 1122 (1961).
9. J. Johnsen, *Arch. Microbiol.* **115**, 271 (1977).
10. Y. Abd-El-Malek, A. Hazem, M. Monib, *Nature* **189**, 775 (1961).
11. Materials and methods are available as supporting material on *Science* Online.
12. C. G. Marshall, I. A. D. Lessard, I. S. Park, G. D. Wright, *Antimicrob. Agents Chemother.* **42**, 2215 (1998).
13. V. M. D'Costa, E. Griffiths, G. D. Wright, *Curr. Opin. Microbiol.* **10**, 481 (2007).
14. S. J. Projan, *Antimicrob. Agents Chemother.* **51**, 1133 (2007).
15. J. L. Parke, D. Gurian-Sherman, *Annu. Rev. Phytopathol.* **39**, 225 (2001).
16. T. Z. DeSantis *et al.*, *Appl. Environ. Microbiol.* **72**, 5069 (2006).
17. J. R. Cole *et al.*, *Nucleic Acids Res.* **35**, D169 (2007).
18. D. L. Wheeler *et al.*, *Nucleic Acids Res.* **28**, 10 (2000).
19. W. Ludwig *et al.*, *Nucleic Acids Res.* **32**, 1363 (2004).

20. We acknowledge the expert assistance of N. Soares and H. Henderson for microbial culturing; P. Girguis and H. White for ARB analysis; D. Bang for HPLC; W. Haas for mass spectrometry; D. Ellison and W. Curtis for sample collection; J.-H. Lee, R. Kolter, and J. Shendure for general discussion; J. Aach for helpful discussions regarding the manuscript; and the U.S. Department of Energy GtL Program, the Harvard Biophysics Program, The Hartmann Foundation, and Det Kongelige Danske Videnskabernes Selskab for funding. 16*S* ribosomal RNA gene sequences were deposited in GenBank with accession numbers EU515334 to EU515623.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5872/100/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S3
References

11 January 2008; accepted 3 March 2008
10.1126/science.1155157

Reversible Compartmentalization of de Novo Purine Biosynthetic Complexes in Living Cells

Songon An,* Ravindra Kumar, Erin D. Sheets,* Stephen J. Benkovic*

Purines are synthesized de novo in 10 chemical steps that are catalyzed by six enzymes in eukaryotes. Studies in vitro have provided little evidence of anticipated protein-protein interactions that would enable substrate channeling and regulation of the metabolic flux. We applied fluorescence microscopy to HeLa cells and discovered that all six enzymes colocalize to form clusters in the cellular cytoplasm. The association and dissociation of these enzyme clusters can be regulated dynamically, by either changing the purine levels of or adding exogenous agents to the culture media. Collectively, the data provide strong evidence for the formation of a multi-enzyme complex, the "purinosome," to carry out de novo purine biosynthesis in cells.

Purines are not only essential building blocks of DNA and RNA, but, as nucleotide derivatives, they also participate in a multitude of pathways in both prokaryotes and eukaryotes (*1*, *2*). Biosynthetically, adenosine and guanosine nucleotides are derived from inosine monophosphate (IMP), which is synthesized from phosphoribosyl pyrophosphate (PRPP) in both the de novo and salvage biosynthetic pathways (Fig. 1). The salvage pathway catalyzes the one-step conversion of hypoxanthine to IMP by hypoxanthine phosphoribosyl transferase (HPRT), whereas the de novo pathway consists of 10 chemical reactions that transform PRPP to IMP. In higher eukaryotes (such as humans), the de novo pathway uses six enzymes, including three multifunctional enzymes: a trifunctional protein, TrifGART, that has glycinamide ribonucleotide (GAR) synthetase (GARS) (step 2), GAR transformylase (GAR Tfase) (step 3), and aminoimidazole ribonucleotide synthetase (AIRS) (step 5) activities; a bifunctional enzyme, PAICS, that has carboxyaminoimidazole ribonucleotide synthase (CAIRS) (step 6) and succinylaminoimidazolecarboxamide ribonucleotide synthetase (SAICARS) (step 7) activities; and a bifunctional enzyme, ATIC, that has aminoimidazolecarboxamide ribonucleotide transformylase (AICAR Tfase) (step 9) and IMP cyclohydrolase (IMPCH) (step 10) activities. In contrast, prokaryotes, such as *Escherichia coli*, use only monofunctional enzymes throughout this pathway, except for the bifunctional ATIC.

Although studies of the individual enzymes in vitro have revealed much about their respective mechanisms of action, a number of in vitro attempts using kinetic analysis and/or binding measurements to demonstrate protein-protein interactions have not been fruitful, with few exceptions (*3*–*6*). In addition, there is scant evidence from in vivo cellular studies for the hypothesis that these enzymes act in concert within a multienzyme complex (*7*). In this study, we used fluorescence microscopy to investigate whether functional multienzyme complexes involved in de novo purine biosynthesis form in living mammalian cells under purine-rich and purine-depleted conditions, which affect the rate of metabolic flux (*8*–*10*).

We selected for study two human (h) enzymes involved with de novo purine biosynthesis as initial candidates for involvement in a multienzyme complex in vivo: the nonsequential hTrifGART protein, which catalyzes steps 2, 3, and 5, and formylglycinamidine ribonucleotide synthase (hFGAMS), which catalyzes step 4. These proteins were fused to either a green fluorescent protein (GFP) or an orange fluorescent protein (OFP) and were transiently expressed in HeLa cells that had been cultured in either purine-rich [minimum essential medium and 10% fetal bovine serum (FBS)] or purine-depleted (RPMI 1640 and dialyzed 5% FBS) media (*9*, *11*). When expressed individually in cells grown in purine-rich media, both the hTrifGART (Fig. 2A) and hFGAMS (Fig. 2C) proteins exhibited diffuse cytoplasmic distributions, as previously found in 293T fibroblast cells and with their prokaryotic counterpart, *E. coli* cells (*7*). However, when these constructs were expressed in purine-depleted cells, we observed cytoplasmic clustering, estimated at ~27% of the hTrifGART-GFP (Fig. 2B) and ~77% of the hFGAMS-GFP (Fig. 2D) proteins. To determine whether the two enzymes colocalized into these clusters in low-purine conditions, we transiently coexpressed hTrifGART-GFP and hFGAMS-OFP in HeLa cells (Fig. 2, G, H, and I); ~60% of the transfected cells exhibited colocalization within the clusters (Fig. 2I).

Z-scan imaging with confocal laser-scanning microscopy confirmed that the hTrifGART-GFP and the hFGAMS-OFP proteins were co-clustered in the cytoplasm (fig. S1). We also observed co-clustering between hFGAMS-GFP and hTrifGART-OFP after we reversed the fusion of the fluorescent probes in HeLa cells that were maintained in purine-depleted media (fig. S2). Additionally, similar cellular localization experiments were carried out with hTrifGART-GFP and hFGAMS-GFP in two additional human cell lines, HTB-125 and HTB-126; the results suggested that the purine-dependent clustering appears to occur in other human cell lines [supporting online material (SOM) Text]. These findings, when combined with the imaging results described

Department of Chemistry, The Pennsylvania State University, University Park, PA 16802, USA.

*To whom correspondence should be addressed. E-mail: sjb1@psu.edu (S.J.B.); eds11@psu.edu (E.D.S.); sua13@psu.edu (S.A.)

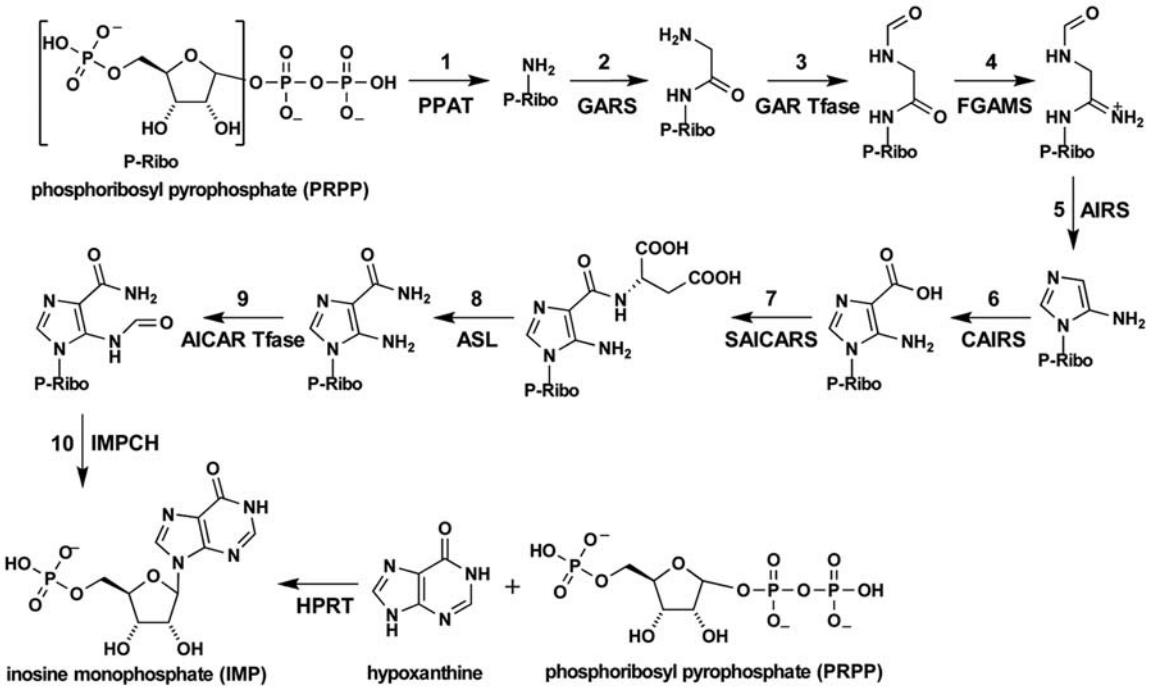

Fig. 1. The de novo purine biosynthetic and salvage pathways. The de novo pathway transforms PRPP to IMP in 10 steps. Steps 2, 3, and 5 are catalyzed by a trifunctional enzyme, TrifGART; steps 6 and 7 are catalyzed by a bifunctional enzyme, PAICS; and Steps 9 and 10 are catalyzed by a bifunctional enzyme, ATIC.

for HeLa cells, suggest that the formation of this multienzyme complex involved in the de novo purine biosynthetic pathway may be functionally relevant in mammalian cells, particularly those challenged with minimal media.

To verify that the clustering we observed was not due to an artifact of expressing the fluorescent fusion constructs but was in fact present with endogenously expressed protein, we conducted immunofluorescence imaging of endogenous hTrifGART protein in HeLa cells grown in purine-depleted media; clustering was clearly observed (Fig. 2E), similar to what we found for the hTrifGART-GFP protein under the same conditions (Fig. 2B). We also identified some clustering of endogenous hTrifGART in the presence of purine-rich media (fig. S3), which we attribute either to the sensitivity of immunofluorescence in detecting protein expression or to the loss of cytosolic materials during fixation. We further demonstrated that endogenous hTrifGART co-clustered with hTrifGART-GFP in HeLa cells that were maintained in purine-depleted media (Fig. 2F), which supports our initial observations with the GFP fusion constructs. Taken together, these results suggest that multienzyme complexes form to satisfy the cellular demand for de novo purine biosynthesis.

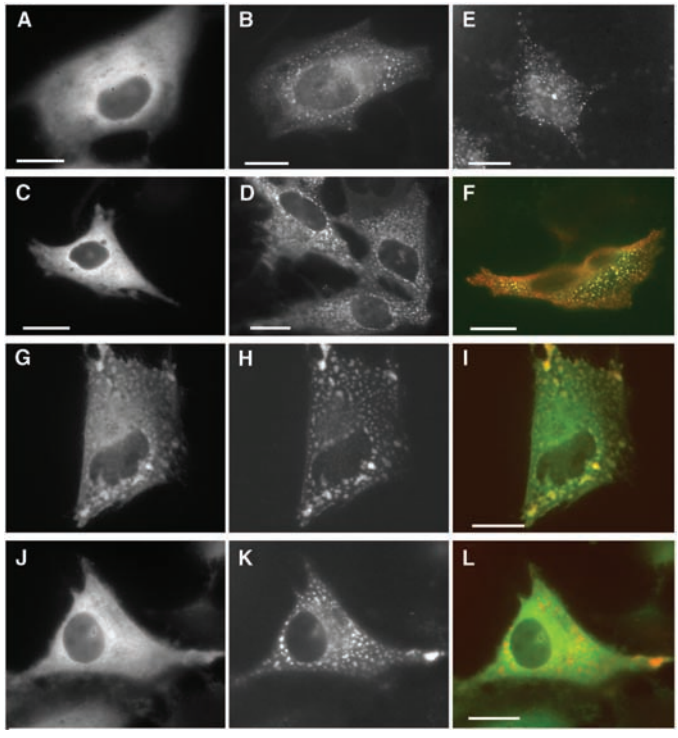

Fig. 2. Cellular localization of enzymes participating in the de novo purine biosynthetic pathway. (**A** to **D**) hTrifGART-GFP [(A) and (B)] and hFGAMS-GFP [(C) and (D)] transiently expressed in HeLa cells in purine-rich [(A) and (C)] and purine-depleted [(B) and (D)] media. (**E**) Immunofluorescence of endogenous hTrifGART protein expressed in HeLa cells cultured under low-purine conditions. (**F**) Merged image by immunofluorescence of hTrifGART-GFP (green) and endogenous hTrifGART (red) expressed in HeLa cells grown in purine-depleted media. (**G** to **I**) hTrifGART-GFP [(G) and green in (I)] co-clusters with hFGAMS-OFP [(H) and red in (I)] in HeLa cells under purine-depleted conditions. (**J** to **L**) The control hC1THF-GFP [(J) and green in (L)], which does not participate directly in de novo purine biosynthesis, does not form clusters or colocalize with hFGAMS-OFP [(K) and red in (L)] in HeLa cells grown in purine-depleted media. Scale bar, 10 μm.

We expanded our studies to include all six human enzymes of the pathway (Fig. 1): PRPP amidotransferase (hPPAT) (step 1), hTrifGART, hFGAMS, hPAICS, adenylosuccinate lyase (hASL) (step 8), and hATIC. As a control, we also included a tetrahydrofolate (H_4F)–utilizing trifunctional enzyme, hC1THF, that possesses *5,10*-methylene-H_4F dehydrogenase, *5,10*-methenyl-H_4F cyclohydrolase, and *10*-formyl-H_4F synthetase activities. hC1THF is not directly involved in the de novo purine pathway but is responsible for synthesizing a key cofactor, *10*-formyl-H_4F, for both hTrifGART and hATIC activities (*5*, *12*). We found that all six enzymes that directly participate in the de novo purine biosynthetic pathway co-clustered with hFGAMS-OFP in the cytoplasm of cells maintained in purine-depleted media (Fig. 2, G to I, and Fig. 3); however, hC1THF-GFP was diffuse throughout the cytoplasm, and thus did not co-cluster with hFGAMS-OFP (Fig. 2, J to L).

When the individual fluorescent protein constructs were transiently expressed in purine-depleted HeLa cells, the frequency of clustering was generally very low: ~5 to 8% for hPPAT-GFP, hPAICS-GFP, hASL-GFP, and GFP-hATIC, but ~27% for hTrifGART-GFP and ~77% for hFGAMS-GFP. However, when co-transfected with hFGAMS-OFP under these conditions, the

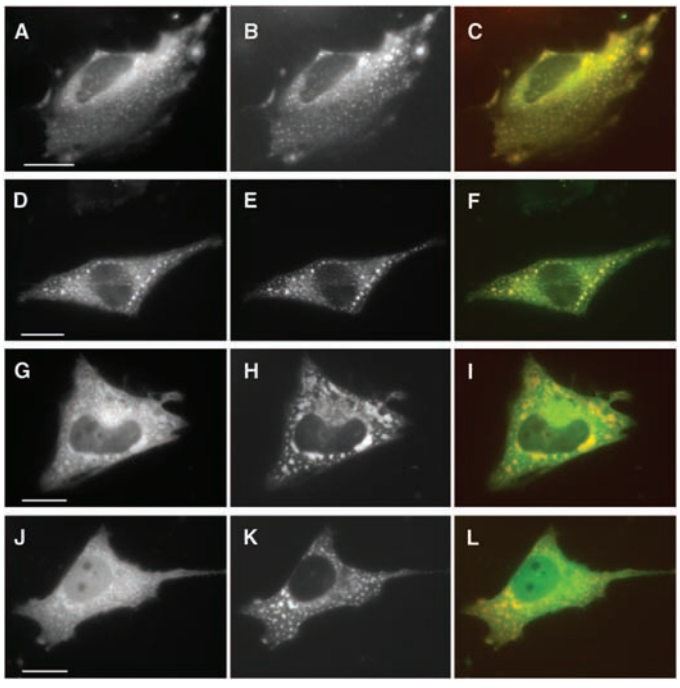

Fig. 3. Colocalization of GFP fusion proteins with hFGAMS-OFP in HeLa cells grown in purine-depleted media. hPPAT-GFP **(A)**, hPAICS-GFP **(D)**, hASL-GFP **(G)**, and GFP-hATIC **(J)** were co-transfected with hFGAMS-OFP (**B**, **E**, **H**, and **K**, respectively). (**C**, **F**, **I**, and **L**) Green channels correspond to the GFP-fusion constructs, and red channels correspond to hFGAMS-OFP in the merged images. Scale bar, 10 μm.

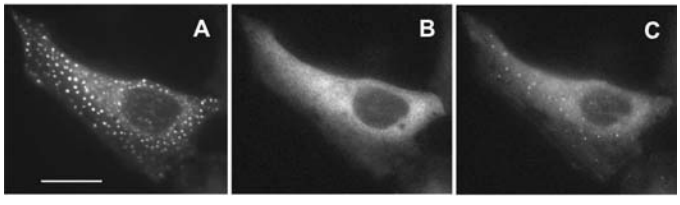

Fig. 4. Reversible formation of clusters by hFGAMS-GFP. (**A**) hFGAMS-GFP clusters when HeLa cells are exposed to purine-depleted medium. (**B**) Clusters disperse within ~0.5 to 1.5 hours upon incubation with purine-rich medium. (**C**) The clusters begin to re-form within ~1 to 2 hours after purine-depleted medium is restored. Scale bar, 10 μm.

clustering frequency increased substantially to ~22 to 32% for hPPAT-GFP, hPAICS-GFP, hASL-GFP, and GFP-hATIC, and ~60% for hTrifGART-GFP. The percentage of colocalization of GFP-fusion proteins with hFGAMS-OFP was calculated by dividing the number of pixels that showed both green and red colors by the total number of pixels that showed either green or red colors (*11*, *13*). Colocalization analysis collectively revealed that 78 ± 6% of GFP-fusion proteins co-clustered with hFGAMS-OFP. However, the analysis for hC1THF-GFP and hFGAMS-OFP exhibited 21 ± 5% colocalization, which suggests that these proteins do not colocalize in clusters: The representative image (Fig. 2L) clearly shows hFGAMS-OFP clusters on a diffuse background of hC1THF-GFP. Scattergram analysis was carried out as an alternative means of examining colocalization, and it showed the compartmentalization of all six enzymes participating in the de novo purine biosynthesis (table S1 and fig. S4). The varying percentage of cells exhibiting clustering under low-purine conditions may be attributed to cells not being synchronized with respect to the cell cycle. The G_1 and S phases require more purines to support RNA and DNA synthesis than the other phases of the cell cycle (*14*); thus, we speculate that the population of the cells exhibiting clustering may represent cells in the G_1 or S phases. This rationale is further supported by the previously measured ~five- to sixfold increase in the level of mRNAs for PPAT and SAICARS in the G_1 and S phases over those in the G_0 phase of synchronized rat 3Y1 fibroblasts (*15*).

To assess the reversibility of the multienzyme complexes, we imaged the cellular distribution of hFGAMS-GFP while purine-depleted medium was exchanged for purine-rich medium, and observed that the hFGAMS-GFP clusters dispersed within ~0.5 to 1.5 hours of HeLa cells being incubated with purine-rich medium (Fig. 4, A and B). When purine-depleted medium was subsequently restored, clusters began to re-form in ~1 to 2 hours (Fig. 4, B and C). Additionally, we performed fluorescence recovery after photobleaching on hFGAMS-GFP in HeLa cells that had been incubated in purine-depleted medium. We clearly observed mobility of the clusters through the rapid fluorescence recovery as a function of time (fig. S5). Collectively, the predominance of clustering appears to be linked to insufficient levels of purines, which suggests that there is a cellular sensor of the external purine conditions and that the mobile clusters represent functional complexes.

Furthermore, we imaged the effects of introducing either an exogenous purine source, hypoxanthine, or a well-known antagonist of hFGAMS, azaserine, to inhibit the metabolic flux in the de novo pathway of HeLa cells grown under purine-depleted conditions (*8*). Hypoxanthine, when added to fibroblast cultures in concentrations as low as 1 to 10 μM, significantly inhibits (>50%) de novo purine biosynthesis through its effect on adenosine and guanosine nucleotides in the cellular pools (*16*, *17*), and the increased levels of adenosine and guanosine nucleotides modulate activities of PPAT as well as PRPP synthetase (*8*). In addition, cells grown in purine-depleted media are deficient in adenosine 5′-triphosphate, with both the restoration of cellular growth rates and the timing of the G_1-to-S transition dependent on both the de novo and salvage pathways (*14*). We propose that the initial increased clustering is a response to hypoxanthine-induced inhibition of the de novo pathway and that the retention of the clusters reflects demand for purines in cellular growth (fig. S6, A, B, and solid squares in E). When we incubated azaserine with purine-depleted cells, we anticipated that hFGAMS clusters would at least be maintained because of the greater demand for purine biosynthesis. An azaserine block increases the rate of the early steps of de novo purine biosynthesis in fibroblasts that lack HPRT activity (*17*), which is consistent with our observed maintenance of the clusters (fig. S6, C, D, and open circles in E). Collectively, the association and dissociation of the multienzyme complex reflected changes in cellular purine levels that were imposed by the addition of external reagents that regulate purine metabolic flux.

From our studies, the various roles that the multienzyme complex plays in cells are not yet clear. These functional complexes may produce efficient substrate channels that link the 10 catalytic active sites. To date, extensive kinetic analyses of *E. coli* PPAT and GARS (*6*), *Acetobacter aceti* purE and purK (bacterial counterparts of CAIRS) (*3*), and recent crystal structures of octameric hPAICS (*18*) have revealed putative substrate channels for individual steps in the de novo pathway. Additionally, clustering of the 10 active sites may provide an efficient means of globally regulating purine flux under varying environmental conditions. However, we have not yet identified a specific subcellular compartment with which these multienzyme complexes associate, in contrast to, for example, the actin cytoskeletal association observed for aldolase (*19*). However, our cross-linking experiments with formaldehyde on living cells grown in purine-depleted media, and subsequent Western blot analysis, revealed that endogenous hTrifGART proteins produced a cross-linked adduct of molecular weight greater than 500 kDa (fig. S7), which suggests that hTrifGART is indeed compartmentalized in close proximity with other proteins, whose identities are under investigation. Moreover, recent proteomic studies suggest that hTrifGART and hFGAMS are probably phosphorylated at binding motifs recognized by 14-3-3 proteins that play important roles in many biological processes, including cell-cycle regulation and signal transduction (*20*), which raises the intriguing possibility that phosphorylation is involved in regulation of the complex.

Although the association of metabolic enzymes has been claimed for enzymes involved in the glycolytic pathway (*21*, *22*), to our knowledge this complex, the glycosome, has not been identified in living mammalian cells. By analogy, the present clusters observed in the de novo purine biosynthetic pathway may constitute a "purinosome." The formation of the purinosome appears to be dynamically regulated by stimulation of de novo purine biosynthesis in response to changes in purine levels. The purinosome may be a general phenomenon in all cell types during specific stages of the cell cycle, along with posttranslation modifications. Because of the relevance of de novo purine biosynthesis to human diseases, the purinosome may represent a new pharmacological opportunity for therapeutic intervention.

References and Notes

1. C. Stasolla, R. Katahira, T. A. Thorpe, H. Ashihara, *J. Plant Physiol.* **160**, 1271 (2003).
2. R. Zrenner, M. Stitt, U. Sonnewald, R. Boldt, *Annu. Rev. Plant Biol.* **57**, 805 (2006).
3. C. Z. Constantine *et al.*, *Biochemistry* **45**, 8193 (2006).
4. C. A. Caperelli, P. A. Benkovic, G. Chettur, S. J. Benkovic, *J. Biol. Chem.* **255**, 1885 (1980).
5. G. K. Smith, W. T. Mueller, G. F. Wasserman, W. D. Taylor, S. J. Benkovic, *Biochemistry* **19**, 4313 (1980).
6. J. Rudolph, J. Stubbe, *Biochemistry* **34**, 2241 (1995).
7. L. T. Gooljarsingh, J. Ramcharan, S. Gilroy, S. J. Benkovic, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6565 (2001).
8. M. A. Becker, M. Kim, *J. Biol. Chem.* **262**, 14531 (1987).
9. T. Yamaoka *et al.*, *J. Biol. Chem.* **272**, 17719 (1997).
10. T. Yamaoka *et al.*, *J. Biol. Chem.* **276**, 21285 (2001).
11. Materials and methods are available as supporting material on *Science* Online.
12. G. F. Wasserman, W. T. Mueller, S. J. Benkovic, W. S. Liao, J. Taylor, *Biochemistry* **23**, 6704 (1984).
13. M. Howarth, K. Takao, Y. Hayashi, A. Y. Ting, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7583 (2005).
14. M. Kondo *et al.*, *J. Biochem. (Tokyo)* **128**, 57 (2000).
15. H. Iwahana *et al.*, *Biochim. Biophys. Acta* **1261**, 369 (1995).
16. L. F. Thompson, R. C. Willis, J. W. Stoop, J. E. Seegmiller, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3722 (1978).
17. M. A. Becker, M. J. Losman, M. Kim, *J. Biol. Chem.* **262**, 5596 (1987).
18. S. X. Li *et al.*, *J. Mol. Biol.* **366**, 1603 (2007).
19. L. Pagliaro, D. L. Taylor, *J. Cell Biol.* **107**, 981 (1988).
20. M. Pozuelo Rubio *et al.*, *Biochem. J.* **379**, 395 (2004).
21. M. E. Campanella, H. Chu, P. S. Low, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2402 (2005).
22. F. R. Opperdoes, P. Borst, *FEBS Lett.* **80**, 360 (1977).

23. We thank M. Kyoung (Pennsylvania State University) for valuable suggestions on data collection, imaging analysis, and statistics; E. Kunze (Pennsylvania State University) for assistance with confocal laser-scanning microscopy; A. A. Heikal (Pennsylvania State University) for sharing HTB-125 and -126 cell lines; and R. Y. Tsien (University of California, San Diego) for providing the pRSET_mOrange plasmid as a gift. This work was supported, in part, by Pennsylvania State University and the Center for Optical Technologies (E.D.S.). Additional acknowledgment is made to the Donors of the American Chemical Society Petroleum Research Fund, for partial support of this research (E.D.S.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5872/103/DC1
Materials and Methods
SOM Text
Figs. S1 to S9
Table S1
References

26 October 2007; accepted 21 February 2008
10.1126/science.1152241

Single-Molecule DNA Sequencing of a Viral Genome

Timothy D. Harris,[1]* Phillip R. Buzby,[1] Hazen Babcock,[1] Eric Beer,[1] Jayson Bowers,[1] Ido Braslavsky,[2] Marie Causey,[1] Jennifer Colonell,[1] James DiMeo,[1] J. William Efcavitch,[1] Eldar Giladi,[1] Jaime Gill,[1] John Healy,[1] Mirna Jarosz,[1] Dan Lapen,[1] Keith Moulton,[1] Stephen R. Quake,[3] Kathleen Steinmann,[1] Edward Thayer,[1] Anastasia Tyurina,[1] Rebecca Ward,[1] Howard Weiss,[1] Zheng Xie[1]

The full promise of human genomics will be realized only when the genomes of thousands of individuals can be sequenced for comparative analysis. A reference sequence enables the use of short read length. We report an amplification-free method for determining the nucleotide sequence of more than 280,000 individual DNA molecules simultaneously. A DNA polymerase adds labeled nucleotides to surface-immobilized primer-template duplexes in stepwise fashion, and the asynchronous growth of individual DNA molecules was monitored by fluorescence imaging. Read lengths of >25 bases and equivalent phred software program quality scores approaching 30 were achieved. We used this method to sequence the M13 virus to an average depth of >150× and with 100% coverage; thus, we resequenced the M13 genome with high-sensitivity mutation detection. This demonstrates a strategy for high-throughput low-cost resequencing.

DNA sequencing and the attendant genetic manipulation it enables have fundamentally altered life science, with the completion of the human genome sequence as a major milestone of this work (*1*, *2*). However, large sample sets—thousands of genomes—are required to analyze many phenomena in which genetics plays a role. With current sequencing technologies, the cost and complexity of such experiments remains limiting (*3*). Having the consensus human genome sequence in hand fundamentally changes the technology requirements for resequencing human genomes. In particular, one can use low-cost techniques with much shorter read lengths and higher parallelism than found with the Sanger capillary electrophoresis methods used to generate the reference genome (*4*).

Several recent reports emphasize the progress in short-read sequencing strategies (*5*–*8*). Although those methods have been used successfully to sequence microbial genomes, their current cost of sequencing, the complexity associated with DNA library preparation, and their use of polymerase chain reaction (PCR) amplification may limit broad application to human genome resequencing. The use of PCR is problematic for three reasons. First, because amplification efficiencies vary as a function of template properties, PCR introduces an uncontrolled bias in template representation. Second, short-read techniques require many more templates than conventional sequencing, and the in vitro manipulations to create libraries with defined sequences at the ends of templates are onerous and expensive in terms of DNA manipulation. Third, errors can be introduced; in a recent large-scale cancer resequencing effort, PCR errors alone accounted for about one-third of initially detected "mutations" (*3*). The fidelity of PCR polymerases is widely reported at 0.5 to 1.0 × 10^{-4} (*9*), a substantial error rate for amplification of single-molecule targets. These limitations can be ameliorated by single-molecule sequencing approaches.

Single-molecule sequencing was proposed as early as 1989 (*10*). Recent work, however, has demonstrated the feasibility of single-molecule sequencing using DNA polymerase to sequence by synthesis (*11*), and a subsequent study of single–RNA polymerase activity shows DNA sequence can be inferred from the serial observation of four identical single-molecule templates (*12*). We have used single-molecule DNA sequencing to resequence the M13 phage genome (*13*). Our sequencing-by-synthesis scheme is diagrammed in Fig. 1. The library preparation process is simple and fast and does not require the use of PCR; it results in single-stranded, poly(dA)-tailed templates. Poly(dT) oligonucleotides are covalently anchored to glass cover slips at random positions. These oligomers are first used to capture the template strands, and then either as a primer for the template-directed primer extension that forms the basis of the sequence reading (Fig. 1) or, optionally, for a template replication step before sequencing (Fig. 2A). Up to 224 sequencing cycles are performed; each cycle consists of adding the polymerase and labeled nucleotide mixture (containing one of the four bases), rinsing, imaging multiple positions, and cleaving the dye labels. For the M13 data reported below, this sequencing process was performed simultaneously on more than 280,000 primer-template duplexes.

This single-molecule sequencing method allows a number of innovations that are not possible with bulk sequencing by synthesis (*5*, *8*). Most of these are related to the principle of asynchronous synthesis; that is, because each template molecule

[1]Helicos BioSciences Corporation, One Kendall Square, Cambridge, MA 02139, USA. [2]Department of Physics and Astronomy, Ohio University, Athens, OH 45701, USA. [3]Department of Bioengineering, Stanford University, and Howard Hughes Medical Institute, Stanford, CA 94305, USA.

*To whom correspondence should be addressed. E-mail: tharris@helicosbio.com