

and general methods for *S. pombe* have been described [S. Moreno, A. Klar, P. Nurse, *Methods Enzymol.* **194**, 795 (1991)]. Unless otherwise indicated, yeast cultures were grown in YES medium at 30°C. YES consists of glucose, yeast extract, and amino acid supplements. HU was used at a concentration of 12 mM. Purification of GST fusion proteins and hexahis-tagged proteins expressed in *S. pombe*, immunoblotting, and kinase assays were performed as described [K. Shiozaki and P. Russell, *Nature* **378**, 739 (1995)].

15. M. Fernandez-Sarabia, C. McInerney, P. Harris, C. Gordon, P. Fantès, *Mol. Gen. Genet.* **238**, 241 (1993).
16. H. Murakami and H. Okayama, *Nature* **374**, 817 (1995).
17. To tag genomic *cds1+* with a sequence encoding two copies of the HA epitope and hexahistidine, we made a Pst I-Not I fragment having nucleotides 205

to 1465 of the *cds1+* open reading frame by PCR. This fragment was introduced into pRIP42-Spc1HA6H after digestion with Pst I and Not I enzymes (22). The construct was linearized with Nhe I and transformed into PR109. Integration and tagging were confirmed by Southern (DNA) hybridization and immunoblotting, with function of *Cds1*^{HAHS} confirmed by lack of HU sensitivity. The *nmt1::GST::cds1+* construct was prepared as described for the *nmt1::GST::chk1+* construct (7).

18. The in vivo interaction between GST:Cds1 and Wee1 shown in Fig. 2D did not require HU treatment, presumably because GST:Cds1 was overexpressed.
19. Cells were grown in YES medium to an A_{600} of 1.0 and then treated with HU (12 mM) for 90 min at 30°C. Synchronous cells were then obtained by centrifugal elutriation with a Beckman JE-5.0 elutriation rotor. These cells were diluted to an A_{600} of 0.3 in YES containing 12 mM HU and grown at 30°C for a fur-

ther 4 hours. Cells were scored for progression through mitosis by microscopic observation. HU sensitivity studies were carried out with cells grown to an A_{600} of 0.3 in YES medium and then diluted to 13,000 cells/ml in YES media containing HU (12 mM), followed by growth at 30°C for a further 8 hours. Samples (75 μ l) were taken at regular intervals, plated onto YES plates, and grown for 4 days at 30°C to determine survival.

20. H. Lindsay *et al.*, *Genes Dev.* **12**, 382 (1998).
21. T. Enoch, A. M. Carr, P. Nurse, *ibid.* **6**, 2035 (1993).
22. K. Shiozaki and P. Russell, *Methods Enzymol.* **283**, 506 (1997).
23. We thank N. Rhind, J.-M. Brondello, C. McGowan, and K. Shiozaki for their advice and the Scripps Cell Cycle Groups for their support and encouragement. Funded by NIH.

29 December 1997; accepted 6 April 1998

Chaos, Persistence, and Evolution of Strain Structure in Antigenically Diverse Infectious Agents

Sunetra Gupta,* Neil Ferguson, Roy Anderson

The effects of selection by host immune responses on transmission dynamics was analyzed in a broad class of antigenically diverse pathogens. Strong selection can cause pathogen populations to stably segregate into discrete strains with nonoverlapping antigenic repertoires. However, over a wide range of intermediate levels of selection, strain structure is unstable, varying in a manner that is either cyclical or chaotic. These results have implications for the interpretation of longitudinal epidemiological data on strain or serotype abundance, design of surveillance strategies, and the assessment of multivalent vaccine trials.

New epidemics of an infectious disease can be triggered by the evolution of a novel antigenic type or strain that evades the acquired immunity within the host population created by its predecessors. The most studied case is the influenza virus, where major shifts in the structure of surface antigens can often trigger worldwide pandemics of the novel variant (1, 2). The antigens that are most likely to exhibit diversity are those under strong selection by host immune responses. These polymorphic antigens may be ranked by the degree to which the associated immune response reduces the reproductive or transmission success of the pathogen. We demonstrated previously that those antigens that elicit the strongest immune response, which in turn have the strongest impact on transmission success, may be organized by immune selection acting within the host population into sets of nonoverlapping variants (3). For example, in the case of two antigens each encoded by a distinct locus with two alleles, namely a

and b at one locus and c and d at the other, four genotypes exist (ac, ad, bc, and bd). One set of nonoverlapping variants is ac and bd (a discordant set) and the other is bc and ad. The pathogen population may exhibit a discrete strain structure where one set of nonoverlapping variants exists at much greater frequency than the other. For this pattern to emerge and be stable over time, the intensity of acquired immunity to a specific variant antigen (encoded by a given allele) within the host population must reduce considerably the transmission success or fitness of all subsequent infections by genotypes possessing that allele. Pathogen populations may therefore be categorized into discrete "strains" or serotypes according to the genetic loci that encode antigens eliciting immune responses with the greatest effect on the transmission success of the infectious agent (3).

Here, we show that pathogens that possess antigens that do not elicit an immune response that is strong enough to induce a discrete stable strain structure may exist as a set of strains that exhibit cyclical or chaotic fluctuations in frequency over time. They may still be organized by immune selection into discrete groups of variants, where all the

members in a given group have different alleles at every locus, but the dominance of the group, relative to that of other groups, may fluctuate widely over time, either cyclically or chaotically. Pathogen antigens that elicit immune responses that have little effect on transmission success will not be organized to express a discrete nonoverlapping strain structure. In this case, all possible allele combinations will be maintained at abundances commensurate with their individual transmission success or fitness.

We define the conditions under which each of these three outcomes arises, in terms of the biological characteristics of both the pathogen and the immune response of the host to the various antigens of the infectious agent, using a model in which a pathogen strain is defined by the m alleles that exist at each of n loci. We take no account of other important biological complications such as mutation, seasonality in transmission, time delays between the acquisition of infection and infectiousness to susceptible hosts, or genetic diversity in the host population, influencing the immune responses to particular antigenic variants. These exclusions are deliberate, because we wish to assess the impact of selection imposed by acquired immunity in the host population on temporal trends in the frequencies of different variants and the evolution of the associated strain structure in the pathogen population.

Strains that do not share any alleles (hereafter referred to as a discordant set) are assumed not to interfere with each others' transmission success or fitness as mediated by host immune responses. The degree to which infection with a given strain limits the ability of another strain that shares any alleles to infect the same host is defined by a cross-protection or cross-immunity parameter γ . If $\gamma = 0$, then the strains do not induce cross-protective responses, whereas if $\gamma = 1$, then there is complete cross-protection. It is also assumed that immunity to a given strain i does not prevent infec-

Wellcome Trust Centre for the Epidemiology of Infectious Disease, Zoology Department, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

*To whom correspondence should be addressed. E-mail: sunetra.gupta@zoology.ox.ac.uk

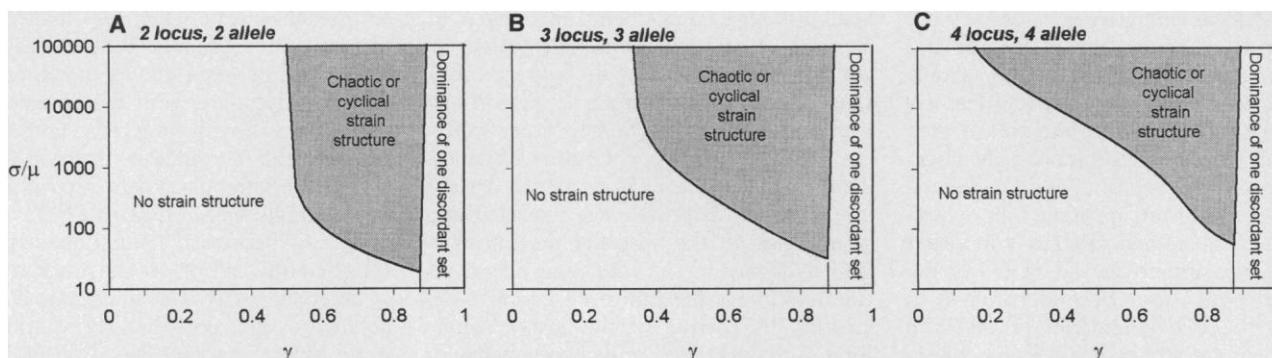


Fig. 1. Strain structure type is shown as a function of the degree of cross-protection (γ) and the ratio of host life-span to pathogen infectious period (σ/μ) for pathogens with (A) two antigenically active loci, each with two alleles; (B) three loci, each with three alleles; and (C) four loci, each with four alleles ($R_0 = 4$ for all strains).

tion by any other strain, but only reduces the probability of transmission of a nondiscordant strain j by a factor $(1 - \gamma)$. This implies that the process of the acquisition of immunity to any one strain is independent to that of its acquisition to any other strain. These simple but realistic biological assumptions can be expressed by a system of differential equations representing the changes in the abundances of each of the m^n strains over time in a genetically homogeneous host population (4).

The model system generated three distinct dynamical behaviors (Fig. 1). (i) When the degree of cross-immunity is below a threshold value, γ_L , all the strains coexist in the host population with stable abundance, and no strain structure (NSS) is apparent. Increasing the degree of cross-protection in this dynamical domain acts to decrease strain abundance because of increased immunological interference between strains. (ii) When the degree of cross-immunity exceeds an upper threshold, γ_U , one discordant set of strains dominates in terms of their prevalence. This situation represents the presence of stable discrete strain structure (DSS). (iii) For $\gamma_L < \gamma < \gamma_U$, no stable strain structure occurs and the relative proportions of the different strains exhibit very complex and often chaotic dynamics with marked fluctuations over time. This is referred to as cyclical or chaotic strain structure (CSS).

Figure 1 plots the parameter regions for which these different dynamical behaviors and associated strain structures pertain. The upper threshold, γ_U , is given by $\gamma_U = 1 - 1/(2R_0)$ if all strains have identical transmission success as defined by the case reproductive rate R_0 [the average number of secondary infections generated by one primary infection in a susceptible host population (5)]. The lower threshold value γ_L is determined by the ratio σ/μ [where $1/\mu$ is host life expectancy and $1/\sigma$ is the average duration of infection (equal to infectious-

ness)], R_0 , and the number of antigen loci and alleles. For pathogens in developed countries, with human life expectancies of around 70 years and infectious periods of between 4 days to 1 year ($1/\sigma = 0.01$ to 1), σ/μ is between 70 to 7000. The large size of the parameter space generating cyclical or chaotic behavior (Fig. 1), for relevant pa-

rameter assignments for the duration of infection and host life expectancy, implies that such complex dynamics may be the norm for many antigenically variable infectious agents that induce moderate cross-protective immune responses in the human host.

Figure 2 portrays the wide range of com-

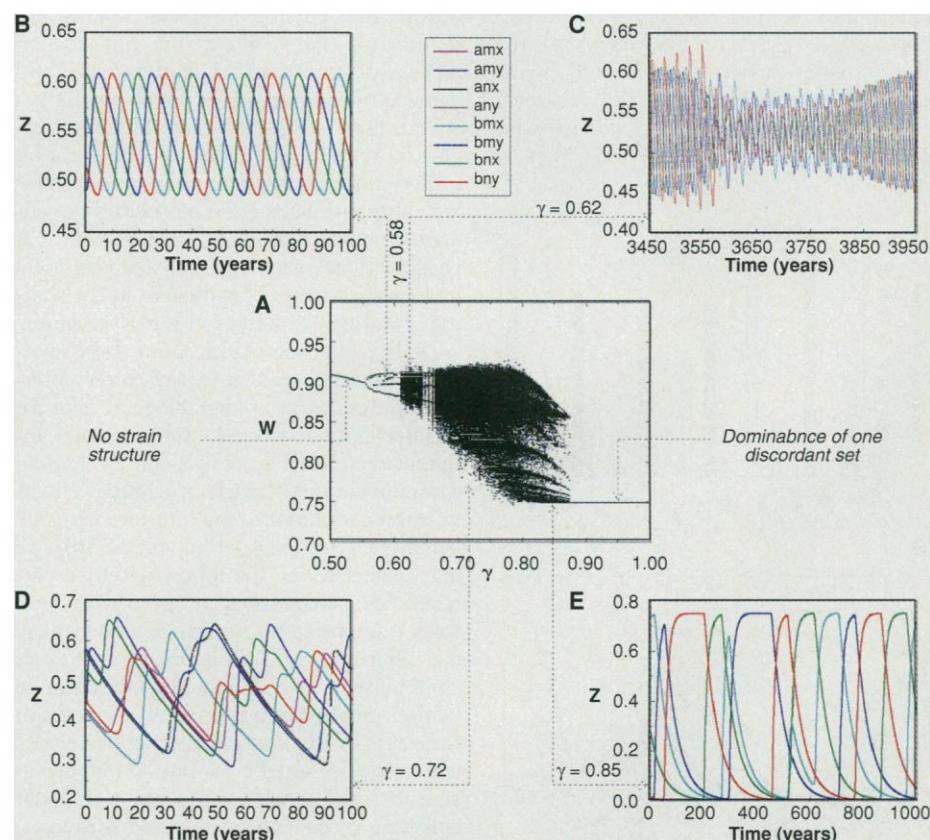


Fig. 2. Long-term population dynamics of pathogens with three loci, each with two alleles. $R_0 = 4$ for all strains, $\sigma = 10$ years $^{-1}$, and $\mu = 0.02$ years $^{-1}$. (A) Bifurcation diagram shows the location of local maxima of w_{amx} (the fraction of the population exposed to strain amx or any strain sharing alleles with amx) as a function of the degree of cross-protection. (B through D) Time-series of the prevalences, z , of each strain in the population: (B) for $\gamma = 0.58$ (nonchaotic limit cycle), (C) for $\gamma = 0.62$ (intermittent chaotic episodes), (D) for $\gamma = 0.72$ (chaotic short-period cycles with decoherence between discordant set members), and (E) $\gamma = 0.85$ (large-amplitude chaos).

plex dynamical behaviors generated by the model in the CSS parameter domain. Initially the strain prevalences follow simple limit cycles (Fig. 2B), but as the value of γ is increased, chaotic intermittency is seen (Fig. 2C), followed by increasingly large-amplitude chaos (Fig. 2, D and E). Several important trends are apparent from extensive numerical studies. First, as γ increases in value, the amplitude and period of the epidemic cycles rise. Second, when γ is close to the two boundaries of the CSS region, complete coherence is seen between the epidemics of strains within discordant sets (that is, the abundances of all strains within a set are identical). This is a direct result of competitive exclusion between genotypes that share alleles where the competition is created by herd immunity. Third, for a large band of cross-protection values in the center of the CSS region, this coherence begins to break down (Fig. 2D), and the trajectories of the prevalence of individual members of discordant sets can be distinguished. The exact dynamical mechanism driving this symmetry breakdown remains unclear, but the behavior appears to be associated with a crossover effect as the system moves from the low- γ region of short-period cycles (a few years) to the large- γ region where long-period (many years) chaos dominates. In the crossover region, both time scales are apparent, with "generation" cycles (with period $\sim 1/\mu =$

host life expectancy) being modulated by shorter period oscillations, the periods of which are determined by the average infectious period ($1/\sigma$) and the transmission success of the pathogen (R_0). In other words, the long periods are determined by host demography and the short periods by the variables that determine the typical course of infection in the host and the transmission dynamics of the infectious agent. As the duration of the infectious period is decreased, the period of the low- γ regime limit cycles decreases, and chaotic behavior increasingly dominates the entire CSS region. Similarly, discordant set decoherence increases as the number of antigenically active loci and alleles increase. Decoherence also increases markedly as the infectious period decreases.

Other theoretical studies indicate that complex dynamics are likely to be common for any pathogen populations that are organized in such a manner that there is cross-immunity within certain strain subsets and none between these subsets. For instance, Andraesen *et al.* (2) modeled influenza by assuming that interstrain cross-immunity exists only between nearest neighbors in a one-dimensional phenotypic space (that is, immunity to strain i affects the transmission success of strains $i - 1$ and $i + 1$) and demonstrated that this form of population structuring can also generate stable limit cycles in systems with four or more strains. Other population-structuring mechanisms may also give interaction matrices that generate complex nonlinear dynamics, but it is particularly relevant that such a fundamental biological property of pathogens as the sharing of antigen variants yields this structure.

A key question arising from these analyses is whether we can expect to see deterministic chaos in epidemiological data for common viral, bacterial, and protozoan antigenically variable pathogens in human communities. Although stochastic effects in small populations may interfere with the persistence of long-period cycles (6), we may expect to see the following broad patterns. For polymorphic antigens that elicit weak immune responses, the abundances of the different strains (as defined by different combinations of alleles) will be determined by their respective transmission successes or fitnesses. In the cyclical or chaotic region, where the degree of cross-protection associated with the antigen is moderate, irregular epidemic cycles will be observed with average periods set by the mean duration of infectiousness in the host (short-period cycles of a few years for infectious periods of a few days to very long periods of many years for infectious periods approaching a month or longer). Concomitantly, there will be a high degree of correlation in the prevalence

or incidence time-series of strains within the same discordant set and a low degree of correlation between strains or serotypes in different discordant sets. For antigens that elicit very strongly protective immune responses, the correlation between strains within the same discordant set will be extremely high, manifesting as DSS.

These theoretical predictions argue for considerable caution in the interpretation of observed patterns in longitudinal epidemiological data sets that are stratified by strain type or serotype based on molecular or immunological taxonomic characteristics. Although the limited time-series data currently available show significant fluctuations in strain incidences for a variety of pathogens (Fig. 3), better quality, longer term data will be required to accurately assess the precise strain structure of a given antigenically variable pathogen population. Essential to this process is the cloning of the relevant antigen genes; for example, the recent cloning of the *Plasmodium falciparum* gene encoding PfEMP-1 (7–9) is crucial to the validation of the hypothesis that PfEMP-1 is the antigen that structures *P. falciparum* into discrete strains (10). The analysis we present suggests that it is also worth monitoring other *P. falciparum* antigens such as the merozoite surface proteins (MSP), which may not have as significant a role in protective immunity as PfEMP-1, because these may be more likely to exhibit cyclical or chaotic fluctuations. A recent study in Senegal (11) involving the typing of successive clinical malaria isolates during a 4-month period of intense transmission shows a large degree of genetic diversity in MSP1 and MSP2 combinations. However, long-term large-scale studies are required to elucidate the precise dynamics of these allele combinations. For bacteria, such as *Neisseria meningitidis*, large-scale nucleotide sequencing studies of genes encoding surface antigens are essential to understand the population structure of these organisms (12, 13) and, in particular, to elucidate whether the strong association observed between VR1 and VR2 epitopes of the PorA protein (2) forms the basis for the fluctuations in lineages of *N. meningitidis* (14). Serotypes of group A streptococci also exhibit rapid fluctuations (Fig. 3) (15). These serotypes are defined by an antiphagocytic cell-surface molecule known as M protein, which appears to occur in nonoverlapping combinations with another variable element known as the serum opacity factor (16), and thus, may exemplify a two-locus multiallele system in a state of CSS.

Such information on pathogen population structure is crucial in many different contexts, including the assessment of vaccine trials where the candidate vaccine

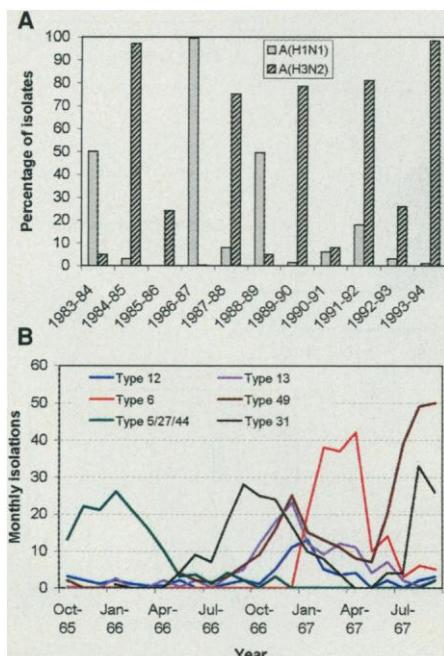


Fig. 3. Examples of dynamical changes in serotype frequencies of two major human pathogens: (A) influenza cases in the United States from 1983–94 (20) and (B) group A streptococcal infections in Minnesota, United States, from 1965–1967 (15).

contains only a subset of the antigens expressed by the pathogen population (17) such as those currently planned for *N. meningitidis* (18) and *Streptococcus pneumoniae* (19). Attributing observed changes in strain structure following mass immunization to the intervention may be problematic, given the complex nonlinear dynamics suggested by our analyses. The intricate behavior of these multistrain systems is a consequence of the selective pressures imposed on the pathogen population by the profile of herd immunity in the host population. This profile is, in turn, conditioned by the prevailing antigenic structure of the pathogen population. It is the subtle interplay between these two factors that leads to the unstable evolutionary dynamics we describe.

REFERENCES AND NOTES

1. A. D. Cliff, P. Haggett, J. K. Ord, *Spatial Aspects of Influenza Epidemics* (Pion, London, 1986).
2. V. Andreasen, J. Lin, S. A. Levin, *J. Math. Biol.* **35**, 825 (1997).
3. S. Gupta *et al.*, *Nature Med.* **2**, 437 (1996).
4. The proportion immune to a given strain i , z_i , is simply given by

$$\frac{dz_i}{dt} = (1 - z_i)\lambda_i - \mu z_i$$

Here, λ_i represents the force of infection of strain i . We assume, for simplicity, that immunity is lifelong. We then define an additional compartment, w_i , which represents those immune to any strain j that shares alleles (at the relevant polymorphic loci) with strain i (including i itself). The dynamics of w_i are then given by (where $j \sim i$ means j shares alleles with i)

$$\frac{dw_i}{dt} = (1 - w_i) \sum_{j \sim i} \lambda_j - \mu w_i$$

Individuals who have never been exposed to any strain sharing alleles with strain i (that is, $1 - w_i$) are completely susceptible to strain i . However, those that have been exposed to a strain sharing alleles with i , but not exposed to strain i itself (that is, $w_i - z_i$), will become infectious with a probability $1 - \gamma$ when they are infected by strain i . With σ being the rate of loss of infectiousness of the host, the dynamics of the proportion of the population infectious for strain i may therefore be represented as

$$\frac{dy_i}{dt} [(1 - w_i) + (1 - \gamma)(w_i - z_i)]\lambda_i - \sigma y_i$$

The impact of genetic exchange on the population structure of infectious disease agents may be examined within this framework by modifying the force of infection term, λ_i , to include the assumption that the progeny of parasites within hosts infectious for two or more strains will consist of defined fractions, Ω_{ijk} , of the various combinations of the different strains, j and k , that may generate strain i through recombination. Because the proportions of infectious hosts are very small, the force of infection of strain i may be approximately represented as $\lambda_i = \beta_i (\bar{v}_i + \sum \Omega_{ijk} y_j y_k)$, where β_i is a combination of parameters affecting the transmission of strain i . The behavior of the model is largely unaffected by the inclusion or precise functional form of the recombination term.

5. R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, Oxford, 1991).
6. B. M. Bolker and B. T. Grenfell, *Philos. Trans. R. Soc. London Ser. B* **348**, 308 (1995); N. M. Ferguson, R. M. May, R. M. Anderson, in *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, D. Tilman and P. Kareiva, Eds. (Princeton Univ. Press, Princeton, NJ, 1997).
- D. I. Baruch *et al.*, *Cell* **82**, 77 (1995).

8. X.-z. Su *et al.*, *ibid.*, p. 89.
9. J. D. Smith *et al.*, *ibid.*, p. 107.
10. S. Gupta, K. Trenholme, R. M. Anderson, K. P. Day, *Science* **263**, 961 (1994).
11. H. Contamin *et al.*, *Am. J. Trop. Med. Hyg.* **54**, 632 (1996).
12. M. C. J. Maiden and I. M. Feavers, in *Population Genetics of Bacteria*, S. Baumberg, P. Young, J. Saunders, E. Wellington, Eds. (Cambridge Univ. Press, Cambridge, 1995), pp. 269–293.
13. I. M. Feavers, A. J. Fox, S. Gray, D. M. Jones, M. C. J. Maiden, *Clin. Diagn. Lab. Immunol.* **3**, 444 (1996).
14. D. A. Caugant *et al.*, *J. Infect. Dis.* **162**, 867 (1990).
15. B. F. Anthony, E. L. Kaplan, L. W. Wannamaker, S. S. Chapman, *Am. J. Epidemiol.* **104**, 652 (1976).
16. J. V. Rakonjac, J. C. Robbins, V. A. Fischetti, *Infect. Immun.* **63**, 622 (1995).
17. S. Gupta, N. M. Ferguson, R. M. Anderson, *Proc. R. Soc. London Ser. B*, in press.
18. P. van der Ley and J. T. Poolman, *Infect. Immun.* **60**, 3156 (1992).
19. D. A. Watson, D. M. Musher, J. Verhoef, *Eur. J. Clin. Microbiol. Infect. Dis.* **14**, 479 (1995).
20. *Morb. Mortal. Wkly. Rep.* **41**, 3 (1992); *ibid.*, p. 5; *ibid.* **42**, 1 (1993); *ibid.* **46**, 1 (1997).
21. S.G. and R.A. thank the Wellcome Trust, and N.F. thanks the Royal Society for grant support. We thank M. Maiden and J. Mathews for their helpful comments.

10 October 1997; accepted 20 March 1998

Identification of Non-Heme Diiron Proteins That Catalyze Triple Bond and Epoxy Group Formation

Michael Lee, Marit Lenman, Antoni Banaś, Maureen Bafor, Surinder Singh, Michael Schweizer, Ralf Nilsson, Conny Liljenberg, Anders Dahlqvist, Per-Olov Gummesson, Staffan Sjö Dahl, Allan Green, Sten Stymne*

Acetylenic bonds are present in more than 600 naturally occurring compounds. Plant enzymes that catalyze the formation of the $\Delta 12$ acetylenic bond in 9-octadecen-12-ynoic acid and the $\Delta 12$ epoxy group in 12,13-epoxy-9-octadecenoic acid were characterized, and two genes, similar in sequence, were cloned. When these complementary DNAs were expressed in *Arabidopsis thaliana*, the content of acetylenic or epoxidated fatty acids in the seeds increased from 0 to 25 or 15 percent, respectively. Both enzymes have characteristics similar to the membrane proteins containing non-heme iron that have histidine-rich motifs.

Over 600 naturally occurring acetylenic compounds are known, many of which occur in higher plants and mosses (1, 2). Knowledge of the enzymatic reactions leading to the formation of a carbon-carbon acetylenic (triple) bond is limited; it is known, however, that in the moss *Ceratodon purpurea*, an acetylenic bond at the $\Delta 6$ position in a C_{18} fatty acid is formed from a carbon-carbon double bond (3), whereas in *Crepis rubra*, oleate is a substrate in the synthesis of 9-octadecen-12-ynoic acid (crepenynic acid) (4).

We have studied the synthesis of acetylenic and epoxy fatty acids in plants of the

genus *Crepis* (2, 5). *Crepis alpina* seed oil is made up of about 70% crepenynic acid, and *Crepis palaestina* seed oil is made up of about 60% vernolic acid (12,13-epoxy-9-octadecenoic acid). Here we characterized the enzymes involved in the biosynthesis of crepenynic and vernolic acid in *C. alpina* and *C. palaestina* (6). Microsomes from developing seeds of *C. alpina*, prepared in the presence of reduced nicotinamide adenine dinucleotide (NADH), converted [^{14}C]linoleate into [^{14}C]crepenynate (Table 1). Similarly, microsomes from developing seeds of *C. palaestina* converted [^{14}C]linoleate into [^{14}C]vernoleate. The results indicated that both the $\Delta 12$ acetylenic bond-forming and $\Delta 12$ epoxidation reactions used acyl chains with a carbon double bond at the $\Delta 12$ position as substrate.

These reactions, and the $\Delta 12$ desaturase, required NADH or NADPH and were inhibited by cyanide (Table 1). Unlike the *Euphorbia lagascae* epoxygenase, which is likely to be a cytochrome P-450-type enzyme (7), the *C. palaestina* epoxygenase was unaffected by carbon monoxide or antibodies to cytochrome P-450 reductase (8). Both the $\Delta 12$ epoxygenation and $\Delta 12$ acetylenation

M. Lee, Svalöv-Weibull AB, S-268 81 Svalöv, Sweden. M. Lenman, A. Banaś, M. Bafor, A. Dahlqvist, P. Gummesson, S. Stymne, Department of Plant Breeding Research, Swedish University of Agricultural Sciences, S-268 31 Svalöv, Sweden. S. Singh and A. Green, Commonwealth Scientific and Industrial Research Organization (CSIRO) Plant Industry, Canberra ACT 2601, Australia. M. Schweizer and S. Sjö Dahl, Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, UK. R. Nilsson and C. Liljenberg, Department of Plant Physiology, University of Gothenburg, S-413 19 Gothenburg, Sweden.

*To whom correspondence should be addressed. E-mail: sten.stymne@vf.slu.se