A self-replicating peptide

David H. Lee, Juan R. Granja, Jose A. Martinez, Kay Severin & M. Reza Ghadiri

Departments of Chemistry and Molecular Biology and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA

THE production of amino acids and their condensation to polypeptides under plausibly prebiotic conditions have long been known[1,2]. But despite the central importance of molecular self-replication in the origin of life, the feasibility of peptide self-replication has not been established experimentally[3–6]. Here we report an example of a self-replicating peptide. We show that a 32-residue α-helical peptide based on the leucine-zipper domain of the yeast transcription factor GCN4 can act autocatalytically in templating its own synthesis by accelerating the thioester-promoted amide-bond condensation of 15- and 17-residue fragments in neutral, dilute aqueous solutions. The self-replication process displays parabolic growth pattern with the initial rates of product formation correlating with the square-root of initial template concentration.

An autocatalytic process in which the reaction product serves as the specific catalyst for its own synthesis by recognizing and promoting coupling of reactant(s) is the most basic form of molecular self-replication (Fig. 1). A few examples of non-enzymatic template-directed autocatalytic systems have been described in which the primary molecular recognition event is based on nucleic-acid-like base-pairing interactions[7–16]. Unlike the nucleic acid bases, which provide a clear framework for establishing complementary molecular interactions, information transfer in polypeptides is inherently more complex and dependent on both the structure and the polypeptide sequence. The present design of the peptide replicator is based on one of the simplest forms of natural protein tertiary structure: namely the α-helical coiled coils. These structures are distinguished by their amphiphilic primary sequence of heptad repeats $(abcdefg)_n$. The first (a) and fourth (d) amino-acid residues make up the hydrophobic core which defines the complementary interhelical recognition surface. Recent studies have established that the identity of the amino acids in these positions and the specificity in the side-chain packing interactions in the hydrophobic core are important determinants of the structures and the aggregation state of coiled coils[17,18]. Conceptually, a given α-helical subunit of a coiled-coil structure can be viewed as a template acting cooperatively to organize other participating peptide subunit(s). We considered that if a similar templating effect could also operate on shorter complementary peptide fragments, it may constitute the basis of an autocatalytic cycle for peptide self-replication.

The self-replication process described is based on a 32-residue peptide similar in sequence to the leucine-zipper homodimerization domain of the yeast transcription factor GCN4 (Fig. 2). The sequence differs from the natural sequence in six mutations out of the 32 residues, five of which are away from the recognition surface: two tyrosine residues were placed on the solvent-exposed surface to provide spectroscopic handles, and alanine and cysteine residues were placed at the ligation site (see below), also on the solvent-exposed surface. The only change made in the interhelical hydrophobic recognition surface was the replacement of asparagine with valine (N16V) to allow the possibility of autocatalysis though one- and/or two-stranded α-helical template structure(s)[17]. The ligation site was chosen to lie on the solvent exposed surface of the α-helical structure to avoid interference with the hydrophobic recognition surface. The amide-bond-forming process was based on the thioester-promoted peptide fragment condensation strategy of Kent[19] in order to circumvent potential side reactions which may arise from reactions at amino and carboxylic acid side-chain functionalities. The reaction was made up of an equal mixture of the electrophilic fragment (preactivated as the thiobenzyl ester COSBn) and the nucleophilic fragment, in the absence or presence of various amounts of template (Fig. 2). The coupling reaction was found to be remarkably regio- and chemo-selective, giving rise to less than 15% side products during the course of the reaction, most of which are due to the slow hydrolysis of the thioester fragment. The reaction products were identified by either direct isolation from reaction mixtures followed by characterization and sequencing by mass spectrometry, and/or by comparison with authentic samples using HPLC co-injections (Fig. 3).

Autocatalysis in template production is unequivocally established when reaction mixtures differing only in the initial concentration of the template are studied (Fig. 4*a*). Marked increases in the initial rates of product (template) formation are observed as the initial concentration of the template is increased. The increase in the initial rates of product formation correlates with the square root of the initial template concentration (Fig. 4*b*). The observed square-root rate profile, which has also been noted in previously characterized synthetic replicators, presumably reflects catalyst (template) inhibition through aggregation[7–13]. The experimental data sets (three sets of five parallel reactions) were also analysed according to empirical rate equations of von Kiedrowski using the program SimFit[16]. These analyses indicate an apparent autocatalytic rate constant $k_a = 29.4 \pm 0.8\ M^{-3/2}\ s^{-1}$ and the background rate constant $k_b = 0.063 \pm 0.008\ M^{-1}\ s^{-1}$ giving an autocatalytic efficiency ($\varepsilon = k_a/k_b$) of approximately 500 (ref. 20). We note that in efficient autocatalytic processes, product formation is expected to display a sigmoidal growth pattern, especially for reactions in which no template is initially present[20]. In this study, sigmoidal growth of template is also noticeable (Fig. 4*a*). As a control when similar reactions were performed in the

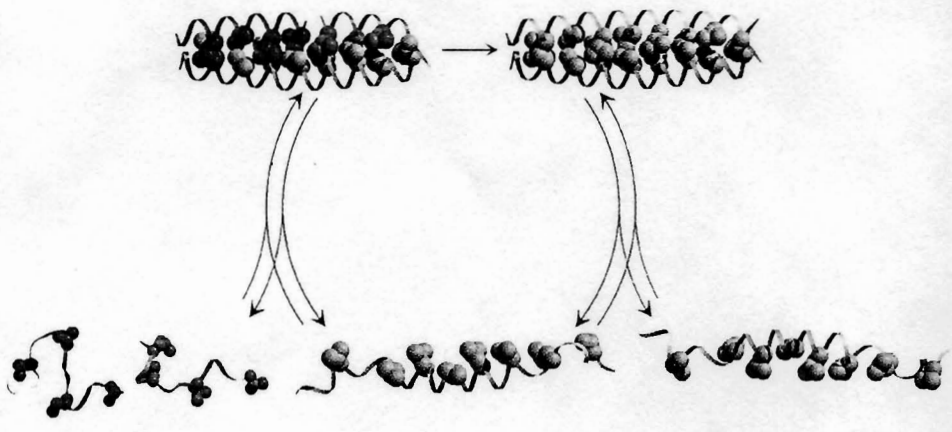

FIG. 1 Diagram of the minimal autocatalytic reaction cycle in the self-replicating α-helical peptide. It consists of two reactive peptide fragments, the electrophile (blue) and the nucleophile (red), each identical in sequence to their putative complementary recognition sites on the template chain (grey). Pre-organization of the reactants on the template, which is mediated by specific inter-helical hydrophobic interactions, promotes peptide fragment condensation. The product, which is an identical copy of the template itself, then becomes part of the autocatalytic cycle to further promote peptide fragment condensation. The single-stranded peptides are depicted as partially unfolded or random coil structures.

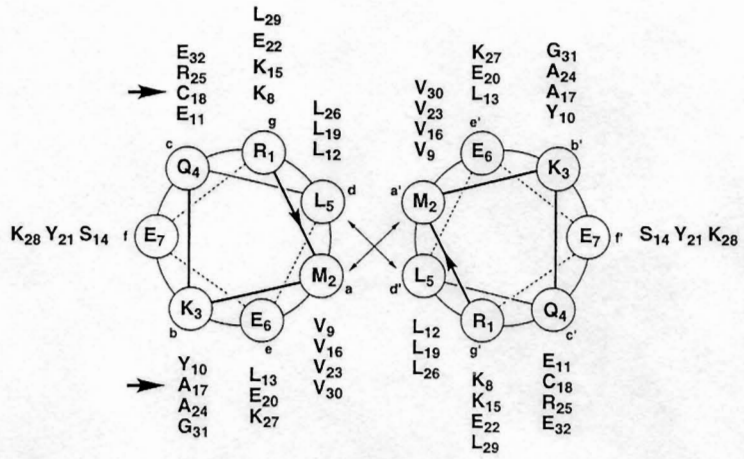

FIG. 2 Helical wheel diagram of the template peptide in the dimeric α-helical coiled-coil configuration emphasizing the heptad-repeat motif (top) and the peptide sequences employed in this study (bottom). The interhelical recognition surface is dominated by hydrophobic packing interactions (positions a and d) and electrostatic interactions (positions e and g). Placement of a single charged polar residue at positions a or d destabilizes the coiled-coil structure by disrupting helix-packing interactions (design of the two crippled templates are based on this principle). Amino acids at positions b, c and f lie on the solvent-exposed surface of the helical structure and do not participate in the molecular recognition processes (arrows indicate the position of the cysteine residue on the nucleophilic peptide fragment and the thioester-activated carboxy terminus of the electrophilic peptide fragment). In the course of the reaction, the electrophilic fragment undergoes an attack by the amino-terminal sulphydryl functionality of the nucleophilic fragment to give the corresponding cysteine thioester intermediate, which rapidly rearranges to give the thermodynamically favoured amide bond at the ligation site. Tyrosine residues were incorporated on the solvent-exposed surfaces to allow accurate determination of peptide concentrations. Furthermore, the template and the electrophilic peptide fragments were acylated at the N termini with 4-acetamidobenzoic acid (Ar) to allow sensitive monitoring of product formation by high-performance liquid chromatography (HPLC).

Template:	ArCONH—RMKQLEEKVYELLSKVA-CLEYEVARLKKLVGE—$CONH_2$
Crippled template 1:	ArCONH—RMKQLEEKVYELLSKVA-CLEYEVAREKKLVGE—$CONH_2$
Crippled template 2:	ArCONH—RMKQEEEKVYELLSKVA-CLEYEVARLKKLVGE—$CONH_2$
Electrophilic fragment:	ArCONH—RMKQLEEKVYELLSKVA—COSBn
Nucleophilic fragment:	H_2N—CLEYEVARLKKLVGE—$CONH_2$

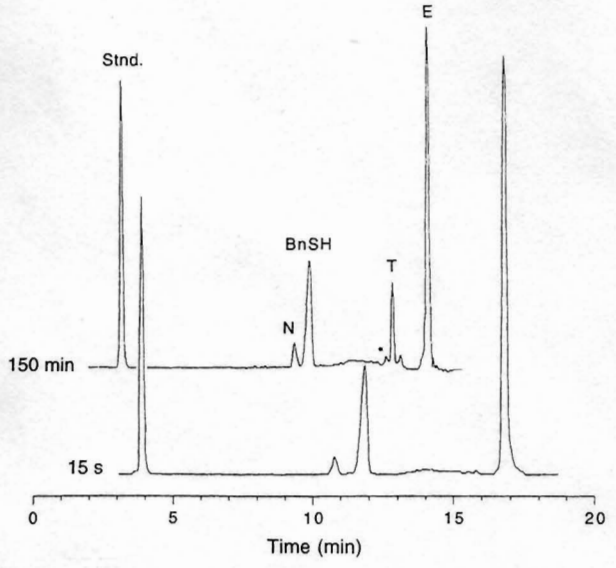

FIG. 3 Reverse-phase HPLC reaction profiles, monitored at a wavelengths of 270 nm, 15 s (bottom trace) and 150 min (top trace) after the initiation of the reaction (* denotes the hydrolysis product, N denotes the nucleophile, T the template, E the electrophile and BnSH is benzylmercaptan).
METHODS. All reactions were started by adding benzylmercaptan-saturated MOPS buffer (final concentration: 154 mM, pH 7.50) to an acidic solution containing equivalent amounts of electrophilic and nucleophilic peptides (final concentration: 90 μM) to give a total volume of 500 μl. Temperature was maintained at 21 °C. Samples (50 μl) were quenched with 2% aqueous trifluoroacetic acid (70 μl) and analysed by reverse-phase HPLC at 270 nm. Concentrations were determined relative to 4-acetamidobenzoic acid as the internal standard (stnd).

presence of guanidinium hydrochloride, which prevents structural organization, no template catalysis was observed (Fig. 4*c*).

A priori, there are at least four pathways that could lead to product formation—three of them constitute the autocatalytic channel (Fig. 5). The ternary complex (Fig. 5*a*) where both reactants are recognized by the template is defined as the 'true' self-replication intermediate. The binary complexes (Fig. 5*b*, *c*) are also plausible autocatalytic intermediates but do not mediate self-replication as only one of the reactants is recognized by the template. To probe the contribution, if any, of these two binary complexes to the overall autocatalytic process, two 'crippled' templates[21,22] were designed, synthesized and characterized (Fig. 2). With respect to the normal template, each of the crippled templates carries a single glutamic acid residue mutation in its hydrophobic recognition surface which severely disrupts interhelical recognition and association. The crippled templates **1** and **2** (Fig. 2) were designed to express the disruptive mutation at the recognition surfaces where the putative binding sites of the nucleophilic and the electrophilic fragments reside, respectively. Therefore, when reactions are performed in the presence of added crippled templates **1** or **2**, any contribution to the overall rates of product formation can only come from the intermediary of the binary complexes shown in Fig. 5*b* or *c*, respectively. When reaction mixtures differing only in the initial concentration of the crippled template were studied—identical reaction conditions to the ones described for the normal template—no increase in the initial rates of product (template) formation was observed (Fig. 4*c*). These control experiments strongly suggest that the contribution of the autocatalytic pathways (shown in Fig. 5*b*, *c*) to the overall product formation is insignificant, and the observed autocatalysis must in fact proceed through the self-replicative ternary intermediate (Fig. 5*a*) in agreement with the described rate analysis and reaction modelling. In addition, preliminary competition studies between the normal and Leu26Ala mutant (in which an alanine residue is substituted for a leucine at position 26) nucleophilic peptide fragments, performed in the presence and/or absence of normal as well as the corresponding mutant templates, have indicated that even such a conservative mutation

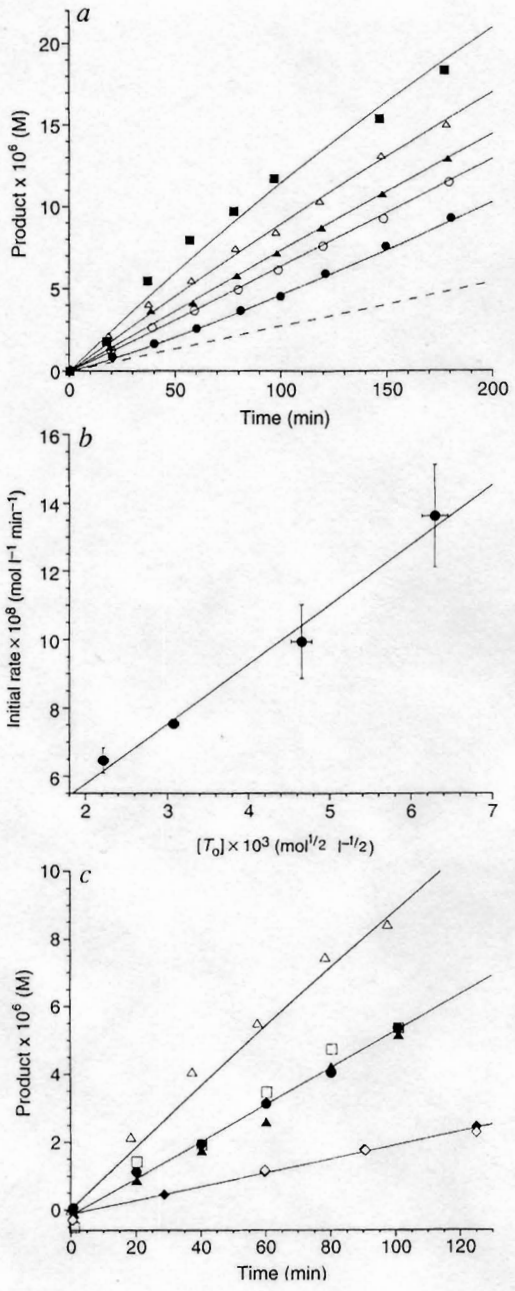

FIG. 4 *a*, Template production as a function of time for reaction mixtures containing different initial concentrations of template. (filled circles) In the absence of any added template, and in the presence of (open circles) 5 μM; (filled triangles) 10 μM; (Δ) 20 μM; and (filled squares) 40 μM initial template concentration. Curves were generated by nonlinear least-squares fit of the data to the empirical rate equation of von Kiedrowski using the program SimFit. The rate equation $dT/dt = (N_0 - T)(E_0 - T)[(k_a(T + T_0)^{0.5} + k_b)]$ describes the initial rate of template formation in two terms: a template-independent term (background) with the apparent rate constant k_b and a template-dependent term (autocatalytic) with a rate constant k_a. N_0, E_0, and T_0 are the initial concentrations of the nucleophile, electrophile and the template, respectively. All reactions were repeated three times (one set is shown) and simulated separately. The dashed line represents the calculated production of template in the absence of autocatalysis. *b*, Initial rate of template production versus the square root of the initial template concentration. Error bars show standard deviations of three independent runs. *c*, Template production as a function of time in the absence of any added template (filled circles); in the presence of 20 μM normal template (open triangles); 20 μM crippled template **1** (open squares), and 20 μM crippled template **2** (filled triangles). The template production is also shown for reactions carried out in pH 7.5 buffered solution of 4.3 M guanidinium hydrochloride in the absence (open diamonds) and the presence of 20 μM normal template (filled diamonds).

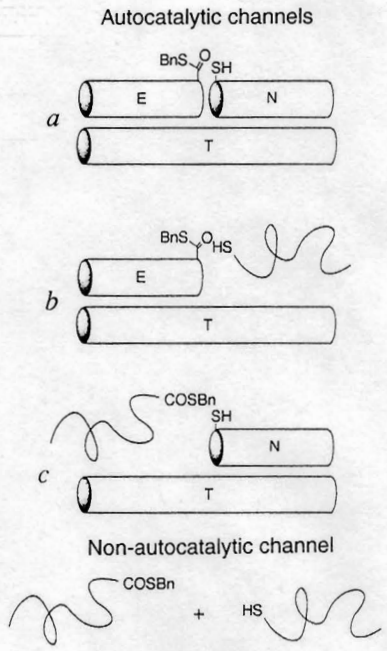

FIG. 5 Schematic representation of various plausible pathways in template production. The autocatalytic channel can be composed of three pathways: reactions can proceed through the intermediary of the ternary complex (*a*) and/or binary complexes (*b* and *c*). These two binary complexes have been shown not to be involved in the present study (see text). Mediation through intermediates of higher aggregation states is however plausible but seems to give a negligible contribution under the reaction conditions employed, based on the kinetic analyses and reaction modellings (see text). The non-autocatalytic channel (background) is probably dominated by a bimolecular event, although unimolecular pathways proceeding through pre-association of the electrophilic and nucleophilic peptide segments is also plausible. However, no noticeable aggregation of the nucleophilic and electrophilic fragments could be detected by circular dichroism spectroscopy even at approximately twofold higher concentrations than used here. Asterisks represent the approximate location of glutamic acid mutations on the helix-recognition surfaces of the crippled templates.

can abolish the self-replication process. Together, these experiments suggest implicitly that sequence-selective molecular recognition on complementary surfaces has a pivotal role in peptide self-replication.

The data presented here establish the chemical feasibility of peptide self-replication. Although the system described is based on α-helical coiled-coil structures, peptide self-replication through the intermediary of larger helical aggregates as well as β-sheet-containing motifs are also likely[23,24]. Furthermore, on the basis of the large number of natural enzymes that are made up of multiple domains, and of recent discoveries indicating that domain swapping is a functional mechanism for interconversion between monomers and oligomeric forms (ref. 25 and refs therein), we suggest the possibility of protein self-replication in which the catalytic activity of the enzyme could also be conserved. Given that amino acids and polypeptide species could have been produced under prebiotic conditions, the possibility of self-replicating peptide sequences should be considered in the early evolution of living systems. □

Received 1 May; accepted 27 June 1996.

1. Miller, S. L. & Orgel, L. E. *The Origins of Life on the Earth* (Prentice Hall, Englewood Cliffs, NJ, 1974).
2. Fox, S. W. & Dose, K. *Molecular Evolution and the Origin of Life* (Dekker, New York, 1977).
3. Orgel, L. E. *Nature* **358,** 203–209 (1992).
4. Oparin, A. I. *Adv. Enzymol.* **27,** 347–380 (1965).

5. Dyson, F. J. *J. Molec. Evol.* **18,** 344–350 (1982).
6. Kauffman, S. A. *J. theor. Biol.* **119,** 1–24 (1986).
7. von Kiedrowski, G. *Angew. Chem. int. Edn. engl.* **25,** 932–935 (1986).
8. Zielinski, W. S. & Orgel, L. E. *Nature* **327,** 346–347 (1987).
9. von Kiedrowski, G., Wlotzka, B. & Helbing, J. *Angew. Chem. int. Edn. engl.* **28,** 1235–1237 (1989).
10. Tjivikua, T., Ballester, P. & Rebek, J. Jr *J. Am. chem. Soc.* **112,** 1249–1250 (1990).
11. Nowick, J. S., Feng, Q., Tjivikua, T., Ballester, P. & Rebek, J. Jr *J. Am. chem. Soc.* **113,** 8831–8839 (1991).
12. von Kiedrowski, G., Wlotzka, B., Helbing, J., Matzen, M. & Jordan, S. *Angew Chem. int. Edn. engl.* **30,** 423–426 (1991).
13. Winter, E. A., Conn, M. M. & Rebek, J. Jr *J. Am. chem. Soc.* **116,** 8877–8884 (1994).
14. Li, T. & Nicolaou, K. C. *Nature* **369,** 218–221 (1994).
15. Pitsch, S. *et al. Helv. chim. Acta* **78,** 1621–1635 (1995).
16. Sievers, D. & von Kiedrowski, G. *Nature* **369,** 221–224 (1994).
17. Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. *Science* **262,** 1401–1407 (1993).
18. Zhou, N. E., Kay, C. M. & Hodges, R. S. *Biochemistry* **31,** 5739–5746 (1992).
19. Dawson, P. E., Muir, T. W., Clark-Lewsi, I. & Kent, S. B. H. *Science* **266,** 776–779 (1994).
20. von Kiedrowski, G. *Bioorg. Chem. Front.* **3,** 113–146 (1993).
21. Wintner, E. A., Tsao, B. & Rebek, J. Jr *J. org. Chem.* **60,** 7997–8001 (1995).
22. Hu, J. C., O'Shea, E. K., Kim, P. S. & Sauer, R. T. *Science* **250,** 1400–1403 (1990).
23. Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E. & Khazanovich, N. *Nature* **366,** 324–327 (1993).
24. Kobayashi, K., Granja, J. R. & Ghadiri, M. R. *Angew. Chem. int. Edn. engl.* **34,** 95–98 (1995).
25. Bennett, M. J., Schlungger, M. P. & Eisenberg, D. *Prot. Sci.* **4,** 2455–2468 (1995).

ACKNOWLEDGEMENTS. We thank S. Kent, P. Dawson and M. Churchill for assistance in peptide synthesis and discussions regarding the peptide ligation strategy, and G. von Kiedrowski for discussion and the generous gift of the program SimFit; the MRC of Canada for a predoctoral fellowship (D.H.L.); NSCORT for a summer visiting scholarship (J.R.G.); the Ministry of Education and Science of Spain for a postdoctoral fellowship (J.A.M.); and Deutsche Forschungsgemeinschaft for postdoctoral fellowship (K.S.). M.R.G. held an Eli Lilly grant from 1994–96.

CORRESPONDENCE should be addressed to M.R.G. (e-mail: ghadiri@scripps.edu).

Trace gas emissions on geological faults as indicators of underground nuclear testing

C. R. Carrigan*, R. A. Heinle*, G. B. Hudson†, J. J. Nitao* & J. J. Zucca‡

* Geosciences and Environmental Technologies Department (L-206), † Isotope Sciences Division (L-231) and ‡ Geophysics and Global Security Department (L-205), Lawrence Livermore National Laboratory, PO Box 808, Livermore, California 94550, USA

UNDERGROUND **nuclear explosions produce trace amounts of distinctive but ephemeral radionuclide gases. In the context of monitoring a comprehensive test ban treaty, the detection of these gases within the territory of a signatory, during a challenge inspection[1–5], may indicate the occurrence of a clandestine nuclear event. Here we report the results of an experiment simulating a well-contained underground nuclear explosion, undertaken to test the ability of natural gas-transport processes to move highly dilute and rapidly decaying radionuclides to the surface. We find that trace gases are transported to the surface within periods of weeks to a year, by flow along faults and fractures driven by barometric pressure variations. Both our observations and related simulations exhibit a chromatographic behaviour, with gases of higher atomic mass and lower diffusivity reaching the surface more rapidly. For a 1-kilotonne nuclear test under conditions identical to those of our experiment, we predict that short-lived ^{133}Xe and ^{37}Ar would be detectable, respectively, about 50 and 80 days after the detonation. Our results indicate that radionuclide sampling along natural faults and fractures, as a forensic tool, can be an extremely sensitive way to detect nearby underground nuclear explosions that do not fracture the surface.**

Sited at a depth of 400 m in the bedded tuffs[6] of Rainier Mesa (Fig. 1*a*) at the Nevada Test Site, a simulated 1-kilotonne (1-kt) nuclear explosion was produced by the detonation of 1.3 million kilograms of chemical explosives in a mined cavity on 22 September 1993. At this depth, the overburden is more than adequate to contain such an explosion. Whereas a somewhat shallower detonation might have produced a collapse crater or extensive fractures connecting the cavity with the surface, this explosion produced no visible structural changes including fissuring at the surface[7]. Such an event represents an extreme challenge for the application of on-site forensic techniques which are intended to provide evidence of the violation of a nuclear treaty. Of these techniques, the sampling of soil-gases for rare, explosion-produced rationuclides at the surface near a suspected underground test is particularly attractive because it is minimally invasive and yields relatively unambiguous results about the occurrence of an event. Our experiment was designed to evaluate the ability of transport processes to carry dilute and rapidly decaying radionuclides from the rubblized detonation site to the surface within the limited period of detectability of these gases.

A gas bottle containing 1.3 m^3 (STP) of ^{3}He was placed in the cavity before loading the explosives. Another bottle with 8 m^3 (STP) of sulphur hexafluoride (SF_6) was positioned just behind a metal bulkhead at the entry to the cavity to minimize any thermal degradation of the tracer although the gas is somewhat resistant to thermal breakdown[8]. Following detonation and release of the tracers, gas chromatography was used to detect the heavy SF_6 gas (relative molecular mass 146) in air samples which have a low background level of less than 3 parts per trillion by volume (p.p.t.v.) at the test site. Similarly, the low atmospheric background level (7.34 p.p.t.v.) of ^{3}He (atomic mass 3) permits a very low threshold for detection using mass spectrometry.

Almost 200 gas samples were obtained over approximately 500 days during which a wide variety of barometric events were recorded by a weather station on Rainier Mesa (elevation 2,286 m above sea level). Most of the sampling occurred in the autumn and spring, when large barometric depressions occur.