velocity (~26 km s^{-1}) and the speed at which the bow shock propagates in a perpendicular direction (~10 km s^{-1}). The recombining solar wind plasma and suprathermal H^0 may also be seen in the wake region. The most likely source of detection of the solar wake would be in Lyα absorption, because the H I Lyα line becomes optically thick at N (H I) ~ 10^{13} cm^{-2} for local cloud properties. A H^0 population heated in the solar system may contribute H I Lyα absorption with an anomalous velocity dispersion toward background stars such as Sirius (located close to the antiapex of solar motion, offset about 4° downwind).

20. For instance, doubling the H^0 density from 0.1 to 0.2 cm^{-3} yields t_e = 3500 to 5200 years ago for the cloud vector (V = −19.0 km s^{-1}, ℓ = 334.5°, b = −1.8°).
21. W. B. Landsman, R. C. Henry, H. W. Moos, J. L. Linsky, *Astrophys. J.* **285**, 801 (1984).
22. A. C. Cummings, and E. C. Stone, Contributions to the 20th International Cosmic Ray Conference, Moscow, U Supl. Ser. R (1987); Contributions to the 21st International Cosmic Ray Conference, Adelaide, South Australia (1990). Because the O and C errors are correlated, they are reduced by 10% (A. C. Cummings, personal communication).
23. P. C. Frisch, D. E. Welty, D. G. York, J. R. Fowler, *Astrophys. J.* **357**, 514 (1990); J. M. Shull and M. van Steenberg, *Astrophys. J. Suppl. Ser.* **48**, 95 (1982). The recombination rate of C^+ and an electron is given by $\alpha_r(C^+) = 4.70 \times 10^{-13}\ T_4^{-0.624}$ $cm^3\ s^{-1}$, where T_4 is the temperature in units of 10^4 K.
24. J. R. Jokipii, in *COSPAR Colloquia Series No. 1, Physics of the Outer Heliosphere*, S. Grzedzielski and E. Page, Eds. (Pergamon, New York, 1990), p. 169.
25. D. C. Morton, *Astrophys. J. Suppl. Ser.* **77**, 119 (1991).
26. J. M. Ajello *et al.*, *Astrophys. J.* **317**, 964 (1987); J. Murthy, *et al.*, *ibid.* **315**, 675 (1987); P. C. Frisch, *COSPAR Colloquia Series No. 1, Physics of the Outer Heliosphere*, S. Grzedzielski and E. Page, Eds. (Pergamon, New York, 1990), pp. 19–21; R. Lallement, *ibid.*, pp. 49–60.
27. J. H. Black and A. Dalgarno, *Astrophys. J. Suppl. Ser.* **34**, 405 (1977).
28. This temperature range is based on Ulysses spacecraft observations of the arrival angle of helium [T = 6700 ± 1500 K (*15*)] and GHRS data for interstellar gas within 1 to 12 pc (see Table 1).
29. The total recombination rate for $Mg^+ \rightarrow Mg^0$ includes radiative and dielectronic contributions. For this comparison, the magnesium-based n_e is derived from the dielectronic recombination rate for $Mg^+ \rightarrow Mg^0$ for the temperature range $1500 < T_e < 60{,}000$ K (including spin-orbit coupling, which mixes autoionizing and nonautoionizing states at low temperatures [H. Nussbaumer, and P. J. Storey, *Astron. Astrophys. Suppl. Ser.* **64**, 545 (1986)]. At T_e ~ 6000 K the radiative (*23*) and dielectronic recombination rates for $Mg^+ \rightarrow Mg^0$ are approximately equal, and so the uncertainty introduced by the $Mg^+ \rightarrow Mg^0$ recombination rate may be small. However, for 6500 K < T < 10,000 K, this uncertainty can be significant. In the calculations of this electron density, Lallement *et al.* (*15*) used a parameterization of a dielectronic recombination by V. L. Jacobs, J. Davis, J. E. Rogerson, M. Blaha [*Astrophys. J.* **230**, 627 (1979)]. However, the lower limit of the effective range of the calculation of Jacobs *et al.* is 10,000 K, so it is inapplicable at surrounding cloud temperatures. In warm ionized nebulae, charge exchange between Mg^0 and H^+ contributes to the formation of Mg^+. This rate is calculated at $\alpha_{ce}(T_4) = 1.74 \times 10^{-9} \exp(-2.210/T_4)$ $cm^3\ s^{-1}$, where $T_4 = T_e/10^4$, in the range $0.8 < T_4 < 2.0$ [R. J. Allan, R. E. S. Clegg, A. S. Dickinson, D. R. Flower, *Mon. Not. R. Astron. Soc.* **235**, 1245 (1988)].
30. D. Pequignot and S. M. V. Aldrovandi, *Astron. Astrophys.* **161**, 169 (1986); F. C. Bruhweiler and Y. Kondo, *Astrophys. J.* **260** L91 (1982).
31. This magnetic field strength represents the total azimuthal field derived for the ordered magnetic field component in the vicinity of the sun. See R. J. Rand and S. R. Kulkarni [*Astrophys. J.* **343**, 760 (1989)] and R. J. Rand and A. G. Lyne (in preparation).
32. V. B. Baranov, M. G. Lebedev, M. S. Ruderman, *Astrophys. Space Sci.* **66**, 441 (1979); V. B. Baranov, M. G. Lebedev, Ju. G. Malama, *Astrophys. J.* **375**, 347 (1991); R. Lallement *et al.*, *ibid.* **396**, 696 (1992).
33. J. D. Slavin, *Astrophys. J.* **346**, 718 (1989); K.-P. Cheng and F. C. Bruhweiler, *ibid.* **364**, 573 (1990).
34. R. Genova *et al.*, *ibid.* **355**, 150 (1990).
35. D. G. York, *ibid.* **204**, 750 (1976).
36. J. V. Vallerga *et al.*, *ibid.* **411**, 729 (1993).
37. I acknowledge the support of National Aeronautics and Space Administration Space Physics Division grant NAGW-2610 during the preparation of this paper. I especially thank R. Lallement and J. Linsky for sending me preprints in advance of publication.

9 February 1994; accepted 14 July 1994

Chemical Sequence Control of β-Sheet Assembly in Macromolecular Crystals of Periodic Polypeptides

Mark T. Krejchi, Edward D. T. Atkins, Alan J. Waddon, Maurille J. Fournier, Thomas L. Mason, David A. Tirrell*

A family of uniform periodic polypeptides has been prepared by bacterial expression of the corresponding artificial genes, with the objective of exploring the potential for control of supramolecular organization in genetically engineered protein-based polymeric materials. The repeating units of the polypeptides consist of oligomeric alanylglycine sequences interspersed with glutamic acid residues inserted at intervals of 8 to 14 amino acids. Crystallization of such materials from formic acid produces β-sheet structures in the solid state, as shown by vibrational spectroscopy, nuclear magnetic resonance spectroscopy, and wide-angle x-ray diffraction. The diffraction results, together with observations from electron microscopy, are consistent with the formation of needle-shaped lamellar crystals whose thickness is controlled by the periodicity of the primary sequence. These results can be used to control solid-state structure in macromolecular materials.

The design and synthesis of artificial proteins is an emerging area of research with important implications for structural biology, materials science, and biomedical engineering. Significant progress has been reported in the design of simple proteins that adopt predictable secondary structures, and preliminary successes in higher order protein folding have been achieved in several laboratories (*1*).

We are interested in controlling the solid-state structures—and in particular the crystal structures—of artificial proteins. Unlike the products of conventional polymerization processes, artificial proteins can be engineered with virtually absolute control of chain length, sequence, and stereochemical purity. Appropriately designed artificial proteins thus represent a new class of macromolecular materials, with properties potentially quite different from those of the synthetic polymers currently available and in widespread use (*2*).

Since the concept of macromolecular structure was established in the early part of this century by Staudinger (*3*), enormous effort has been devoted to the development of polymerization processes that afford improved control of chain architecture and to the elucidation of the relations between molecular structure and supramolecular organization in solid polymers. Keller and others (*4*) in the 1950s established the generality of polymer crystals as folded-chain lamellae, and it is now accepted that the folded-chain architecture is a kinetic trap for essentially any flexible polymer chain. Thus, it is prudent to focus any crystal engineering effort on lamellar structures, with the following key parameters the objects of control: (i) chain conformation, (ii) unit cell structure, (iii) lamellar thickness, and (iv) lamellar surface structure. Toward these ends, we describe herein the design, synthesis, and structural analysis of the family of artificial proteins represented by sequence **1** (Fig. 1).

The design of this simple family of sequences was based on concepts and observations drawn in part from polymer chemistry and physics and in part from structural biology. The repeating alanylglycine (AlaGly) dyads were selected to form extended β strands and to assemble into β-sheet crystal "stems" in the lamellar aggregate. Poly(AlaGly) and a variety of AlaGly-rich polypeptides (including *Bombyx mori* silk fibroin) are known to adopt β-sheet struc-

M. T. Krejchi, A. J. Waddon, D. A. Tirrell, Department of Polymer Science and Engineering, University of Massachusetts, Amherst, MA 01003, USA.
E. D. T. Atkins, H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, United Kingdom.
M. J. Fournier and T. L. Mason, Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, MA 01003, USA.
*To whom correspondence should be addressed.

tures in the solid state (5). The choice of three to six dyads reflects two considerations: (i) the known crystallization behavior of synthetic aliphatic polyamides, which suggests that kinetic factors should restrict the length of the stem to that defined by six to eight lateral hydrogen bonds (6, 7) and (ii) the observation that the egg-stalk protein of *Chrysopa flava* folds with a periodicity corresponding to eight amino acid residues (8). We therefore focused on the repeating octapeptide **1a** and examined **1b** through **1d** in corollary fashion. Finally, glutamic acid was inserted into the sequence for reasons of size, polarity, and reactivity. The large size of Glu relative to Ala and Gly should make inclusion of the acid in the lamellar interior sterically unfavorable. The intersheet spacings in poly(AlaGly), for example, are on the order of 4 to 5 Å (5), whereas salts of poly(Glu) in the β form are characterized by intersheet distances of ~13 Å (9). During the crystal growth process, strong interactions with the solvent or with ions should segregate Glu residues to the surface of the growing aggregate, and the reactivity of the Glu side chain is attractive from the point of view of subsequent chemical modification, either of the isolated chain or of the lamellar assembly. Finally, we chose Glu because it is the weakest β-sheet former of all 20 amino acids in the Chou-Fasman scheme for prediction of protein secondary structure (*10*).

A simple representation of the targeted crystalline structure of protein **1a** is shown in Fig. 2: a folded-chain lamellar aggregate roughly 30 to 35 Å thick, with antiparallel β-sheet crystal stems arranged on an orthorhombic unit cell, and with the periodic Glu residues decorating the lamellar surfaces. The following paragraphs describe the synthesis of **1a** and the evaluation of this structural hypothesis.

The double-stranded oligonucleotide **2** (Fig. 1), which encodes two copies of the octapeptide repeating unit of **1a**, was synthesized with the use of phosphoramidite chemistry (*11*). The design of the coding sequence reflects the following considerations: (i) avoidance of codons rarely used by the host strain, (ii) minimization of strict sequence periodicity, and (iii) elimination of all Ban I restriction sites other than those flanking the target sequence. The coding sequence was multimerized and transferred to the expression vector pET3-b (*12*), and expression of target protein **3** (Fig. 1) was monitored by the incorporation of ^{3}H-glycine after induction during logarithmic phase growth of a culture of *Escherichia coli* strain BL21(DE3)pLysS transformed with the resulting recombinant plasmid. Reduction in the growth rate upon induction was

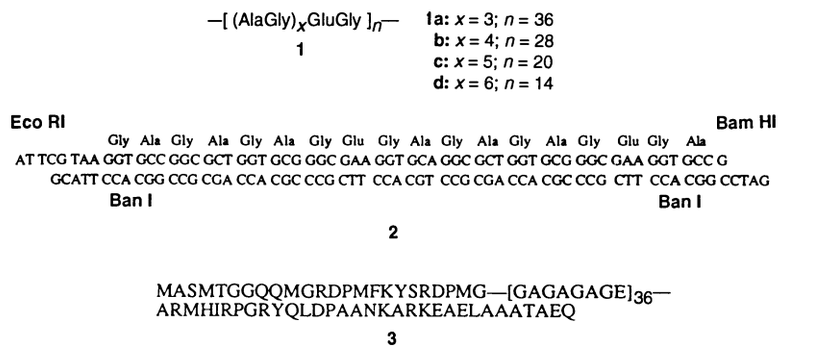

Fig. 1. Basic structure of the family of protein sequences synthesized in this experiment (**1**). Double-stranded oligonucleotide (**2**). Target protein (**3**).

coincident with the accumulation of a new protein product in electrophoretic analyses of whole cell lysates (Fig. 3). This product was not present in untransformed cells, in cells transformed with pET3-b, or in uninduced cells transformed with the target gene. The product migrates as a single band on SDS-polyacrylamide gels with an apparent molecular weight of about 60,000, and, although this is higher than expected (26,741), similar highly acidic proteins migrate anomalously under these conditions (*13*). The target protein was confined to the soluble fraction of the whole cell lysate, and a simple procedure involving sequential pH adjustments with glacial acetic acid afforded a substantial enrichment in the supernatant (Fig. 4). The product was isolated by ethanol precipitation at −10°C, and the periodic portion of **3** was liberated from the flanking sequences by cleavage with cyanogen bromide (*14–16*).

The predominance of the antiparallel (ap) β-sheet structure in the crystalline polypeptide is revealed in the key features of the infrared (IR), Raman, and cross-polarization magic angle spinning nuclear magnetic resonance (CP/MAS NMR) spectra of samples crystallized from 70% formic acid. The IR spectrum of **1a** exhibits amide I, II, and III vibrational modes at 1623, 1521, and 1229 cm^{-1}, respectively (Fig. 5A), characteristic of the β-sheet conformation (*17, 18*), and the weak amide I component at 1698 cm^{-1} indicates the regular alternation of chain direction that defines the apβ-sheet architecture (*19*). Close examination of the amide I region (Fig. 5B) reveals the presence of four vibrational modes unrelated to an apβ-sheet architecture. The weak vibrations centered at 1710 and 1744 cm^{-1} can be attributed to carbonyl stretching of the side chain carboxyl groups of Glu; however, the amide I component observed at 1652 cm^{-1} and the weaker shoulder at 1665 cm^{-1} indicate that some fraction of the polypeptide chain has

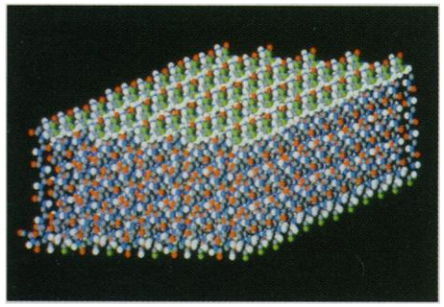

Fig. 2. Computer-generated space-filling molecular model of a hypothetical lamellar crystal constructed from stacked, folded β sheets of **1a**.

adopted a secondary structure other than that of the apβ-sheet. Normal mode calculations (*20*) suggest that these amide I components (as well as the prominent shoulder at 1550 cm^{-1} in the amide II region) may arise from reverse turns of either the β or γ type.

The Raman spectrum of **1a** (Fig. 5C) exhibits the amide I band at 1664 cm^{-1} and the splitting of the amide III band into two components at 1260 and 1228 cm^{-1} characteristic of the apβ-sheet conformation (*18, 21*). The broad, weak amide II band observed at 1535 cm^{-1} is expected for the apβ-sheet structure (*22*), and, as in the IR spectrum, it is plausible to associate the bands in the 1300 to 1330 cm^{-1} region with turn structures (*20*).

The CP/MAS ^{13}C NMR spectrum of **1a** was obtained at 50.6 MHz; the observed chemical shifts are compared with those of poly(L-AlaGly) in Table 1. All of these assignments are consistent with an apβ-sheet architecture (*23*); however, close examination of the line shape of the β-carbon signal of Ala reveals the presence of a shoulder at 16.8 ppm. The chemical shift of this shoulder is close to that observed for the β-carbon in Ala in the silk-I crystal structure (*24*), but the absence of the corresponding carbonyl peak at 176 ppm suggests that this signal is not associated with the silk-I structure. The intensity of this signal was not reduced by washing the sample with formic acid, a good solvent for

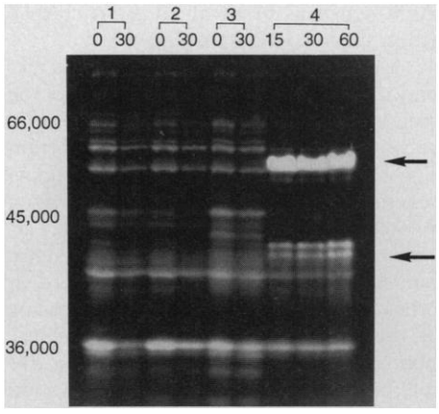

Fig. 3. Electrophoretic analysis of protein expression by in vivo labeling of proteins derived from: (1) BL21(DE3) pLysS, (2) BL21(DE3) pET3-b pLysS, and (3) BL21(DE3) pLysS pET3bMK36 (no induction), and (4) BL21(DE3) pLysS pET3bMK36 at various time points (in minutes) after addition of IPTG. The expression of **3** was monitored by the incorporation of ^{3}H-glycine into the target protein. Frozen primary stock, or single colonies from YT plates containing BL21(DE3) pLysS pET3bMK36, BL21(DE3) pLysS pKF526 (positive control), BL21(DE3) pLysS pET3-b (no insert control), and BL21(DE3) pLysS (no plasmid control) were used to inoculate 5 ml of fresh M9AA (*39*) without glycine. Incubation with appropriate antibiotic selection was carried out at 37°C for 12 hours. Portions of these cultures (90 to 100 μl, 1.455 optical density units per milliliter) were used to inoculate fresh M9AA medium (10 ml) with selection. These cultures were incubated with vigorous agitation at 37°C until the [absorbance at 600 nm (A_{600})] reached 0.4 to 0.5. Then ^{3}H-glycine (50 μCi, 1 mCi/ml) was added, and the cells were incubated at 37°C for 15 min. IPTG (15 μl, 0.2 M) and an additional volume of ^{3}H-glycine (100 μl, 1 mCi/ml) were added. At several time points during cell growth, aliquots (1 ml) were removed and the cells were harvested by centrifugation at 13,500*g* for 45 s. The cells were washed with fresh media, resuspended in 50 μl of 10 mM tris (pH 7.5) and 10% glycerol and lysed by the addition of 50 ml of 100 mM tris (pH 7.5), 4% SDS, 4 mM EDTA, 2% β-mercaptoethanol, and 20% glycerol. The lysates (20 to 30 ml, 1.007 optical density units per milliliter) were heated at 95°C for 5 to 10 min and electrophoresed at 10 mA for 14 hours in a 10% polyacrylamide gel. Gels were stained with Coomassie brilliant blue R, destained, soaked in autoradiography enhancer (Enlighten), and dried on Whatman 3MM filter paper. Photographic film was exposed to the dried gel at −80°C for 2.5 days.

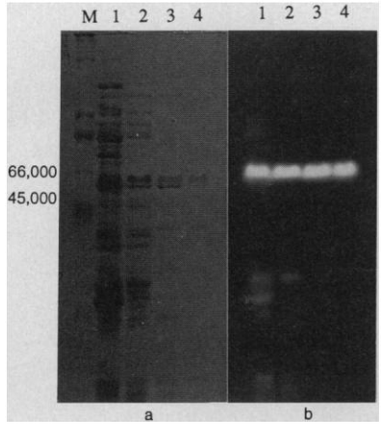

Fig. 4. Electrophoretic analysis of the purification of **3**. Proteins were visualized by (a) Coomassie blue staining and (b) radiolabeling. Lanes M, 1, 2, 3, and 4 are protein molecular weights standards, crude lysate, pH 5.0 supernatant, pH 4.5 supernatant, and pH 4.0 supernatant, respectively. Radiolabeled cells from an induced cell culture (see legend to Fig. 3) were pelleted by centrifugation at 13,000*g* for 2 min, washed with fresh media, resuspended in 1 ml of TE (10 mM tris, 1mM EDTA), placed on ice, and disrupted by sonication at 40 units of power (Branson Sonifier) for 30 s. The lysate was clarified by centrifugation at 13,000*g* for 10 min. The clarified supernatant was acidified to pH 5.0, pH 4.5, and pH 4.0 sequentially, by the addition of aqueous acetic acid (1%); after each pH adjustment, the solution was placed on ice for at least 4 hours and then centrifuged at 13,000*g* for 10 min. At each pH an aliquot (50 μl) of the supernatant was removed, and an equal volume of a solution containing 100 mM tris (pH 7.5), 4% SDS, 4 mM EDTA, 2% β-mercaptoethanol, and 20% glycerol was added. Samples were then analyzed by electrophoresis as described in Fig. 3.

the silk-I modification, reinforcing the conclusion that the shoulder is not associated with a silk-I fraction. It would be consistent with the IR and Raman analysis to propose that this signal arises from turn structures.

Wide-angle x-ray diffraction patterns of crystal mats of **1a** were obtained with the x-ray beam parallel to the plane of the mat. The diffraction patterns exhibit discrete Bragg diffraction signals consistent with crystalline polymer; a representative example is shown in Fig. 6. The wide-angle diffraction signals index on an orthorhombic unit cell with dimensions a = 0.948 nm, b = 1.060 nm, and c (chain axis) = 0.695 nm. The measured and calculated interplanar spacings are listed in Table 2, and the most prominent diffraction signals are indexed in the schematic pattern shown in Fig. 6C.

These unit cell parameters are commensurate with previously published x-ray diffraction results from various silk fibroins (*25*) and synthetic polypeptides (*26*) that exhibit similar crystalline structures. The value of the 0.948-nm spacing was assigned on the basis of its second diffraction order observed at 0.474 nm. This value is close to the characteristic hydrogen-bond distance observed in many nylons (*27*, *28*) and is attributed to the distance between hydrogen-bonded chains in the apβ-sheet structure. Thus, it is possible to associate a with the hydrogen-bond direction. The value of 1.060 nm for b, or more specifically its second diffraction order at 0.53 nm, represents the average intersheet stacking periodicity. Because the amino acid side chains within the apβ sheets decorate the sheet surfaces, the precise value for this spacing is a function of the amino acid composition (*29*). Values for this spacing have been reported as low as 0.34 nm for poly(Gly) (*30*) and as high as 0.79 nm for *Nephila senegalenis* silk fibroin (*25*), a silk that contains a high percentage of amino acid residues with bulky side chains. The values reported for poly(L-AlaGly) (*17*) and the β-form of poly(Ala) (*26*) are 0.44 nm and 0.54 nm, respectively; thus, the spacing observed for **1a** is in line with expectation.

In the x-ray diffraction pattern in Fig. 6A the orientation direction is along a, that is, the hydrogen-bond direction. This result suggests that the orienting forces are acting on crystalline aggregates, not on individual polymer chains. The geometry of these crystalline aggregates must be such that they have a substantially greater dimension in the a direction as compared with the other two mutually orthogonal directions b and c. Transmission electron microscopy (Fig. 7) reveals the presence of needle-shaped polymer crystals similar to those observed for conventional nylons and polypeptides. Fraser *et al.* (*17*) obtained a similar texture by stroking viscous solutions of poly(L-AlaGly), and the same orientation occurs in the thin fibrous egg stalks drawn by the green lace-wing fly *Chryosopa* (*8*). In the texture observed in **1a**, the directionally coincident a and a^* axes lie along the meridian line and the 200 and 400 diffraction signals appear as arcs on that meridian (Fig. 6, A and B). All the $0k0$ and 00ℓ diffraction signals appear on the equator, suggesting cylindrical symmetry about the a, a^* axes. Previous analyses of the apβ-sheet structure suggest that, of the 00ℓ diffraction signals, only the 002 and 006 diffraction signals have significant intensity. In **1a**, the 002 diffraction signal is obstructed because it coincides with 030 (the interplanar spacings differ by only 1.4%); however, the 006 signal is observed at 0.116 nm with a Debye-Scherrer x-ray diffraction camera.

Differences in line broadening of the various diffraction signals which relate in a consistent way to their Miller indices are observed in crystal mats of **1a**. The 200 and 400 arcs are particularly sharp. This demands long-range correlation order along a, consistent with the observation that the crystallites are of substantial length in one dimension, most reasonably assigned as the a direction. Diffraction signals with indices of the type $hk\ell$ are considerably broader than those with indices $hk0$. In particular, we estimate that the 211 diffraction signal is approximately tenfold broader than its 210 near neighbor. These signals are sufficiently close to each other in reciprocal space that we can make a realistic compar-

ison and are not compromised by other signals generated by the lattice. This comparison gives an unmistakable indication that the coherent scattering length in the *c* direction is small (2 to 4 nm) compared with the other crystallographic directions. The relative broadness of the 010 and 020 diffraction signals allows us to establish that the coherent scattering length in the *b* direction, or the intersheet direction, is intermediate between those in the *a* and *c* directions, again consistent with the results of electron microscopy. Analysis of images such as that in Fig. 7 suggests crystal dimensions of ~2000 Å (assigned as *a*) by ~150 Å (assigned as *b*) by about 40 to 100 Å (from shadow lengths; assigned as *c*).

A low-angle x-ray diffraction pattern taken with the beam parallel to the plane of the mat is shown in Fig. 6C. This pattern is composed of a prominent arced diffraction signal focused on the equator corresponding to a spacing of 3.6 ± 0.1 nm, together with a weaker and diffuse order. This low-angle pattern resembles an interparticle interference function and is similar to that observed in other chain-folded systems (*7*, *31*). A spacing of this magnitude oriented along the mat normal is indicative of stacks of lamellae; this interpretation is bolstered by the observation that this diffraction signal disappears after swelling of the mats in glycerol. In contrast, with the exception of a small change in the degree of orientation, the character and spacings observed for all other diffraction signals remain unchanged after swelling. These results indicate that this long period is associated with the dimensions of an intercrystalline repeat, and, although the spacing is not a direct measure of lamellar thickness (*32*), a long period of 3.6 nm is commensurate with the calculated fold-to-fold distance of 2.8 nm (and with the additional volume required by the Glu side chains) in the lamellar structure (Fig. 2).

X-ray diffraction and electron microscopic analyses of polymers **1b** through **1d** have provided important insights into the structure of **1a**. A full description of these studies will be published separately (*33*); we discuss herein only those features relevant to the model proposed below. Wide-angle x-ray diffraction patterns obtained from sedimented crystal mats of **1b** through **1d** index on orthorhombic unit cells, with dimensions listed in Table

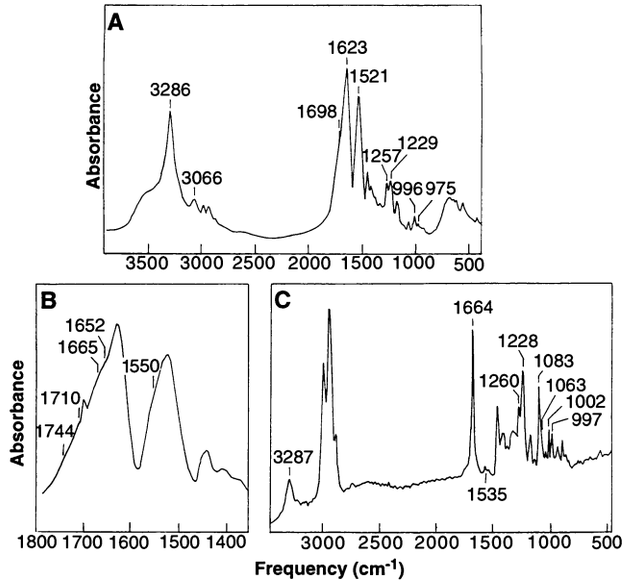

Fig. 5. Vibrational spectra of **1a**. (**A**) Fourier transform infrared spectrum of **1a**. The IR spectra were obtained on an IBM IR32 Fourier transform IR spectrophotometer using KBr pellets prepared from powder samples of **1a** at a concentration of 0.2% (by weight). We prepared samples by stirring cyanogen bromide–cleaved protein (30 mg/ml) in formic acid (70%) overnight at room temperature. The resulting gel was washed with formic acid (70%) and methanol, then resuspended in methanol (40 ml), and incubated at −10°C for 2 days to extract residual formic acid. Powder samples were collected by centrifugation and dried in vacuo overnight at room temperature. (**B**) Expansion of the amide I region of the IR spectrum of **1a**. (**C**) Raman spectrum of **1a**.

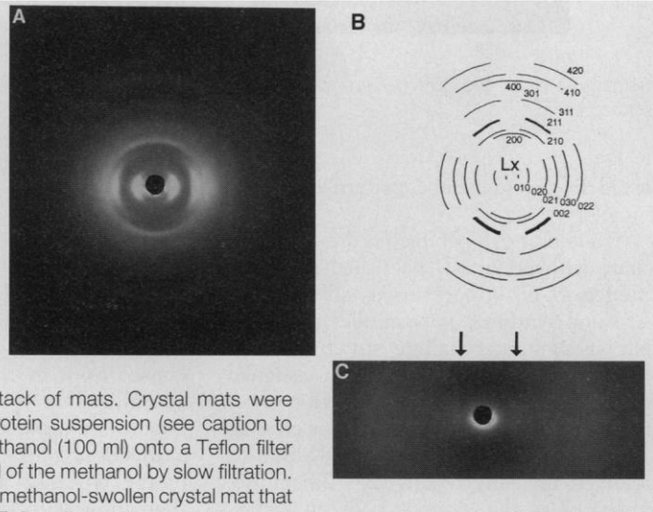

Fig. 6. X-ray diffraction patterns from oriented samples of **1a**. (**A**) Wide-angle x-ray pattern taken with a Statton camera with the beam parallel to the plane of the mat. (**B**) Schematic illustration of the prominent diffraction signals observed with indexing details. The symbol LX refers to the low-angle signal. (**C**) Low-angle diffraction pattern taken with a Statton camera with the beam parallel to the plane of a stack of mats. Crystal mats were obtained by allowing the protein suspension (see caption to Fig. 5) to sediment from methanol (100 ml) onto a Teflon filter (10 μm) followed by removal of the methanol by slow filtration. This procedure resulted in a methanol-swollen crystal mat that was placed between two Teflon filters, then sandwiched between two Whatman filters, and dried overnight at room temperature under compression between two glass plates. X-ray diffraction patterns of sedimented crystal mats were obtained with a Statton-type evacuated x-ray camera. The nickel-filtered CuKα sealed beam source was collimated with a system of 200-μm pinholes and directed parallel to the mat surface. The diffraction patterns were recorded on x-ray film with specimen-to-film distances of 3 cm (wide angle) and 17 cm (low angle), respectively.

Table 1. Chemical shifts and assignments for selected peaks observed in the CP/MAS ^{13}C NMR spectra of **1a** and the β form of poly(L-alanylglycine) (PLAG) (*23*).

Chemical shift (ppm)		Assignment
1a	PLAG	
49.9	48.5	Ala C_α
20.7	20.0	Ala C_β
171.4	171.8	Ala C=0
43.6	43.3	Gly C_α
171.4	168.4	Gly C=0

Table 2. Comparison of observed diffraction signal spacings (d_o) in **1a** with those calculated (d_c) for an orthorhombic unit cell with dimensions a = 0.948 nm, b = 1.060 nm, c = 0.695 nm.

hkℓ	d_o	d_c
010	1.060	1.060
020	0.530	0.530
021	0.424	0.421
030	0.353	0.353
002		0.348
012		0.330
031		0.215
022	0.298	0.291
040		0.265
006	0.116	0.116
070	0.151	0.151
200	0.474	0.474
210	0.437	0.433
211	0.368	0.367
301	0.285	0.288
311		0.278
400	0.241	0.238
410	0.230	0.231
420	0.212	0.216

3. In all cases, the orientation is the same as that observed in **1a**, and morphologies similar to that of Fig. 7 are observed by electron microscopy.

A detailed examination of these x-ray patterns indicates that signals with indices of the type $hk\ell$ sharpen as the length of the repetitive alanylglycine segment increases. Specifically, the signal indexed 211 is approximately twice as sharp in **1b** as in **1a**, and the sharpness increases by another factor of 1.5 from **1b** to **1c**. Upon further increase in the stem length to **1d**, no further change in the character of the 211 signal is noted, an indication that the correlation length along the *c* axis increases as the length of the segment is increased.

In addition to changes in the width of signals indexed $hk\ell$, changes in interplanar spacing and relative intensities of signals indexed 0*k*0 occur on progression through the series. A decrease in the relative intensity of the 010 signal is observed as the number of repeating alanylglycine dyads is increased. This decrease in intensity is accompanied by an overall reduction in the intersheet spacing of 10.2%, with the greatest change (9.4%) occurring between **1a** and **1b**. This result correlates with the decreasing proportion of Glu residues in the chemical sequence and is in agreement with previously published work relating the effect of amino acid composition on the intersheet spacing (*29*). These changes in the character and spacing of the 010 signal are accompanied by an increase in the relative intensity and a corresponding decrease in the interplanar spacing of the 020 signal.

All of the experimental evidence collected on the solid-state structure of **1a** (and other members of series **1**) strongly supports a crystalline apβ-sheet architecture involving a chain-folded lamellar structure as the basic crystalline unit. This structure (Fig. 8) is constructed by the lateral stacking of apβ-sheets. The planes of the sheets are parallel to the longer dimension of the crystalline lamellae and perpendicular to the lamellar surface. The hydrogen-bond direction in the sheet is parallel to the long dimension of the crystalline lamella, and the chain axis is normal to the lamellar surface.

The crystal dimensions observed for all of these polymers indicate that the lamellar thickness is always shorter than the molecular length of the chains. Therefore, the chains must fold back at the lamellar surfaces and reenter the crystalline lamellae, as generally accepted for flexible polymers (*34*). A direct consequence of the chain reversal is the antiparallel arrangement of adjacent polypeptide chains within the lamellar crystal as demanded by the spectroscopic and x-ray analysis of these materials. The long period spacings observed in the low-angle, x-ray patterns, in combination with the selective line broadening observed in the wide-angle patterns are consistent, in all cases, with chain folding predominantly in register with the chemical sequence periodicity as dictated by the necessity to position the Glu residue at the fold surface. However, no direct evidence of surface confinement of Glu has been obtained. The folding is considered to be of the regular adjacent reentry type, but probably with occasional irregular folding (*35*). Signal intensity is believed to be directed to the 010 planes through the pairing of adjacent sheets, perhaps as a consequence of a "polar" arrangement of alanyl methyl groups, as proposed first for *B. mori* silk fibroin (*36*) and subsequently for poly(AlaGly) (*17*). In this arrangement, methyl groups protrude from only one face of the sheet, and it has been suggested that this leads to juxtaposition of faces bearing only glycine protons or only alanyl side chains, respectively (*36*). The resulting alternation in intersheet distances along the *b* axis of the unit cell enhances the 010 intensity.

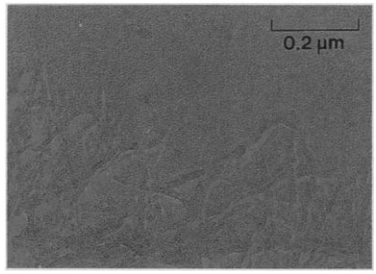

Fig. 7. Transmission electron micrograph of crystalline lamellae of **1a** shadowed with Pt-Pd.

Table 3. Unit cell dimensions of **1a** through **1d**.

Sample	Unit cell dimensions (nm)		
	a	*b*	*c*
1a	0.948	1.060	0.695
1b	0.948	1.028	0.695
1c	0.957	0.970	0.695
1d	0.964	0.962	0.695

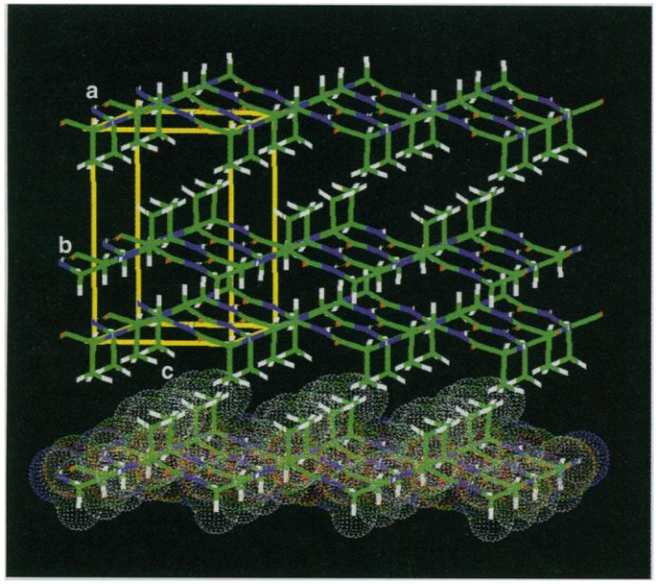

Fig. 8. Computer-generated representation of the solid-state structure of **1a**, as determined by vibrational and NMR spectroscopy and x-ray diffraction. The model comprises folded β sheets stacked along the *b* axis (vertical) of the orthorhombic unit cell. The chain (*c*) axis lies horizontally, and the hydrogen-bonding (*a*) direction extends perpendicular to the plan of the page. A "polar" arrangement of sheets is proposed, with alanyl methyl groups juxtaposed between the top two sheets, and glycyl protons sandwiched between the middle sheets. At the bottom is a spacefilling representation showing the volume requirements of a portion of a single folded sheet.

REFERENCES AND NOTES

1. See, for example, S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, *Science* **262**, 1680 (1993), and references therein.
2. D. A. Tirrell, M. J. Fournier, T. L. Mason, *Curr. Opin. Struct. Biol.* **1**, 638 (1991).
3. H. Staudinger, *Berichte d. D. Chem. Gesellschaft* **53**, 1073 (1920).
4. A. Keller, *Philos. Mag.* **2**, 1171 (1957); P. H. Till, *J. Polym. Sci.* **24**, 301 (1957); E. W. Fischer, *Z. Naturforsch. Teil A* **12**, 753 (1957).
5. R. D. B. Fraser and T. P. MacRae, *Conformation in Fibrous Proteins* (Academic Press, New York, 1973).
6. P. Dreyfuss and A. Keller, *J. Macromol. Sci. Phys. B* **4**, 811 (1970).
7. E. D. T. Atkins, M. Hill, S. K. Hong, A. Keller, S. Organ, *Macromolecules* **25**, 917 (1992).
8. A. J. Geddes, K. D. Parker, E. D. T. Atkins, E. Beighton, *J. Mol. Biol.* **32**, 343 (1968).
9. H. D. Keith *et al.*, *Biopolymers* **7**, 775 (1969).
10. P. Y. Chou and G. D. Fasman, *Biochemistry* **13**, 222 (1974).
11. L. J. McBride and M. H. Caruthers, *Tetrahedron Lett.* **24**, 345 (1983).
12. The oligonucleotide was ligated into the polylinker region of the cloning vector pUC18 [C. Yanisch-Perron, J. Vieira, J. Messing, *Gene* **33**, 103 (1985)], and the resulting recombinant plasmid was cloned in *E. coli* strain DH5aF′. The coding sequence was verified by double-stranded DNA sequencing; Ban I digestion then yielded a 48–base pair fragment encoding two repeats of sequence **1a**. This "monomeric" DNA was used to produce a population of multimers by ligation with T4 DNA ligase, and the multimers were inserted into the unique Ban I site of p937.51, a cloning vector that contains an origin of replication derived from pBR322 and a gene encoding chloramphenicol acetyltransferase [the nonpalindromic cohesive ends generated by Ban I suppress the formation of inverted repeats upon polymerization of the monomeric DNA, thus preserving the integrity of the coding sequence during this step (F. A. Ferrari, C. Richardson, J. Chambers, S. C. Causey, T. J. Pollock, U.S. Patent Appl.

927,258,1986)]. The resulting mixture was used to transform *Escherichia coli* strain HB101 cells, and a plasmid encoding 36 repeats of sequence **1a** was selected for further characterization. The target DNA was subcloned in pET3-b to yield pET3bMK36, an expression vector in which transcription is driven by bacteriophage T7 RNA polymerase (*38*). In this construction, the target protein sequence is flanked by plasmid-derived amino- and carboxyl-terminal extensions of 23 and 32 residues, respectively (sequence **3**). The host used for protein expression was *E. coli* strain BL21(DE3) pLysS, which carries a gene encoding T7 RNA polymerase incorporated into the bacterial chromosome under *lacUV5* control. This configuration allows protein production to be induced by isopropyl β-D-galactopyranoside (IPTG). The ancillary plasmid pLysS provides a low-level source of T7 lysozyme, which inhibits T7 RNA polymerase and suppresses the basal level of protein expression (*38*).

13. K. P. McGrath, M. J. Fournier, T. L. Mason, D. A. Tirrell, *J. Am. Chem. Soc.* **114**, 727 (1992).
14. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
15. B. J. Smith, *Methods in Molecular Biology: New Protein Techniques* (Humana, Totowa, NJ, 1988), p. 70.
16. The structure of the cleaved protein product was confirmed by ^{1}H NMR spectroscopy, amino acid analysis, and combustion analysis. Calculated amino acid analysis (mol percent): Ala, 37.4; Glx, 12.3; Gly, 49.7. Values found: Ala, 35.6; Glx, 13.1; Gly, 49.3. Calculated combustion analysis (percent by weight; including 2.5% H_2O by weight): C, 45.1; H, 6.2; N, 19.2. Values found: C, 46.1; H, 6.2; N, 18.0. Typical isolated yields of **3** corresponded to ~25% of total cellular protein. Thermogravimetric analysis in air revealed an onset of degradation at 235°C. No thermal transitions were evident by differential scanning calorimetry from room temperature to 235°C.
17. R. D. B. Fraser, T. P. MacRae, F. H. C. Stewart, E. Suzuki, *J. Mol. Biol.* **11**, 706 (1965).
18. W. H. Moore and S. Krimm, *Biopolymers* **15**, 2483 (1976).
19. T. Miyazawa and E. R. Blout, *J. Am. Chem. Soc.* **83**, 712 (1961).
20. S. Krimm and J. Bandekar, *Adv. Protein Chem.* **38**, 181 (1968).
21. B. G. Frushour and J. L. Koenig, *Biopolymers* **14**, 2115 (1975).
22. B. G. Frushour, P. C. Painter, J. L. Koenig, *J. Macromol. Sci. Rev. Macromol. Chem. C* **15**, 29 (1976).
23. The CP/MAS ^{13}C NMR spectra of **1a** were obtained at 50.6 MHz on powder samples with the use of a Bruker AC 200 spectrometer equipped with a DOTY solids probe and an IBM solids rack. Measurements were collected at a spinning speed of ~4000 Hz with a 5-μs 90° pulse and a cross polarization time of 2 ms. A line-broadening factor of 50 was applied during data processing. Chemical shifts were compared with those reported by H. Saito *et al.*, *Macromolecules* **17**, 1405 (1984).
24. M. Ishida, T. Asakura, M. Yokoi, H. Saito, *Macromolecules* **23**, 88 (1990).
25. J. O. Warwicker, *J. Mol. Biol.* **2**, 350 (1960).
26. L. Brown and I. F. Trotter, *Trans. Faraday Soc.* **52**, 537 (1956).
27. A. Keller, *J. Polym. Sci.* **36**, 361 (1959); E. D. T. Atkins *et al.*, *J. Polym. Sci. A* **2**, 863 (1972); P. Dreyfuss and A. Keller, *J. Polym. Sci. B* **8**, 253 (1970); J. H. Magill *et al.*, *Polymer* **22**, 43 (1981).
28. G. Hirichsen, *Makromol. Chem.* **166**, 291 (1973).
29. F. Lucas *et al.*, *J. Mol. Biol.* **2**, 339 (1960).
30. G. Nemethy and M. P. Printz, *Macromolecules* **5**, 755 (1972).
31. K. Sakaoku, N. Morosoff, A. Peterlin, *J. Polym. Sci. Polym. Phys. Ed.* **11**, 31 (1973).
32. C. G. Vonk, in *Small Angle X-ray Scattering*, O. Glatter and O. Kratky, Eds. (Academic Press, New York, 1982), p. 443.
33. M. T. Krejchi, E. D. T. Atkins, M. J. Fournier, T. L. Mason, D. A. Tirrell, unpublished results.
34. D. C. Bassett, *Morphology of Polymers* (Cambridge Univ. Press, Cambridge, 1981).
35. We have recently completed a spectroscopic study of the related polypeptide **4**, which provides persuasive direct evidence for adjacent reentry folding

—[(AlaGly)$_3$GluGly(GlyAla)$_3$GluGly]—,

[A. Parkhe, M. J. Fournier, T. L. Mason, D. A. Tirrell, *Macromolecules* **26**, 6691 (1993)].
36. R. E. Marsh, R. B. Corey, L. Pauling, *Biochim. Biophys. Acta* **16**, 1 (1955).
37. F. W. Studier, A. H. Rosenberg, J. J. Dunn, *Methods Enzymol.* **185**, 60 (1990).
38. Media were prepared according to J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 2, 1989), p. A. 3.
39. We thank J. Cappello for the gift of p937.51 and for helpful advice, S. L. Hsu for guidance in vibrational spectroscopy, and L. K. Thompson for help in obtaining solid-state NMR spectra. Y. Deguchi developed methods for the preparation of polymers **1b** through **1d**. We are grateful to A. D. Parkhe for preparing Fig. 2 and to S. Cooper for Fig. 8. Supported by grants from the Polymers and Genetics Programs of the National Science Foundation (NSF) and by the NSF Materials Research Laboratory at the University of Massachusetts.

4 April 1994; accepted 7 July 1994

Stress Triggering of the 1994 M = 6.7 Northridge, California, Earthquake by Its Predecessors

Ross S. Stein, Geoffrey C. P. King, Jian Lin

A model of stress transfer implies that earthquakes in 1933 and 1952 increased the Coulomb stress toward failure at the site of the 1971 San Fernando earthquake. The 1971 earthquake in turn raised stress and produced aftershocks at the site of the 1987 Whittier Narrows and 1994 Northridge ruptures. The Northridge main shock raised stress in areas where its aftershocks and surface faulting occurred. Together, the earthquakes with moment magnitude $\mathbf{M} \geq 6$ near Los Angeles since 1933 have stressed parts of the Oak Ridge, Sierra Madre, Santa Monica Mountains, Elysian Park, and Newport-Inglewood faults by more than 1 bar. Although too small to cause earthquakes, these stress changes can trigger events if the crust is already near failure or advance future earthquake occurrence if it is not.

The 17 January 1994 Northridge earthquake was the most costly shock in the history of the United States, underscoring the vulnerability of urban areas to earthquakes. The event struck on a blind or buried thrust fault (*1*) inclined to the south. The 1971 **M** = 6.7 San Fernando earthquake struck on adjacent thrust faults inclined to the north (*2*). Both earthquakes were responses to crustal compression across the greater Los Angeles area. Not only did aftershocks of the San Fernando (*3*) and Northridge (*4*) earthquakes spatially overlap (Fig. 1), but the 23-year span between the events is small relative to their probable thousand-year repeat times (*5*), suggesting that the two shocks are related. Here we argue that the San Fernando shock increased stress at the future Northridge rupture zone by up to 2 bars, potentially advancing its occurrence by two decades. This hypothesis is supported by the observation that aftershocks of the 1971 and 1994 earthquakes were concentrated where the stresses are calculated to have risen, and aftershocks were sparse where the stresses are calculated to have dropped.

We calculate the Coulomb stress change caused by one earthquake on the rupture surface of a subsequent shock or

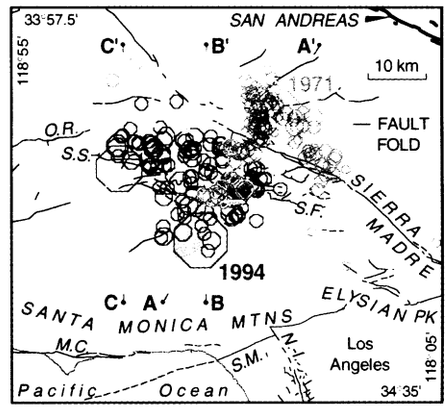

Fig. 1. Overlapping aftershocks of the 1971 San Fernando (blue; first year, $\mathbf{M} \geq 2$) and 1994 Northridge (red; first 24 days, $\mathbf{M} \geq 3$) earthquakes. Sites of mapped secondary surface faulting or cracked ground (green) (*19*): N-CP, Northridge-Canoga Park; GH, Granada Hills; PC, Potrero Canyon; DR, Davidson Ranch. Faults: O.R., Oak Ridge; S.S., Santa Susana; S.F., San Fernando; M.C., Malibu Coast; S.M., Santa Monica; N.-l., Newport-Inglewood. Cross-section orientations of Fig. 3 are also shown.

R. S. Stein, U.S. Geological Survey, Mail Stop 977, Menlo Park, CA 94025, USA. E-mail: stein@andreas.wr.usgs.gov.
G. C. P. King, Institut de Physique du Globe, Strasbourg 67084, France. E-mail: king@klakmuf.u-strasbg.fr.
J. Lin, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. E-mail: jian@galileo.whoi.edu.