

Phosphotransferase sequence homology

SIR—I wish to report a remarkable sequence homology which is found in a number of different enzymes all of which confer antibiotic resistance by acting as phosphotransferases. A sequence with many similar features is also found in a number of different protein kinases.

The homologies were found in the course of developing a programming system for the analysis of sequence patterns. Details will be reported elsewhere, but, briefly, programs are written in a specially-designed string-processing language, for which there is a C interpreter. Databases can be searched for matches to templates composed of elements of different types: thus, X is matched only by X; <XYZ>, by X or Y or Z; <> by anything; and <.XY>, by anything excepting X and Y. Gaps are not allowed. A tolerance on the match can be specified to find not only the perfectly matched sequence but also sequences with one or more errors.

The sequences are shown in the accompanying figure. There are two regions separated by a variable number of amino acids; the first region was used to generate a search template. Originally this was taken from the aminoglycoside phosphotransferases (see figure legend for details) aph.1 to 5 and the template used was H<>D:h<><> <>N:h:h:h, where :h stands for <FLIMVWY> (single letter code), the large hydrophobic amino acids. This template found the viomycin phosphotransferase. The homology between the whole region and that of the aph enzymes is very striking, and particularly so because there is no other sequences resemblance between these enzymes. A similar sequence is found in the hygromycin phosphotransferases, which share no other homologies either with each other or with the other phosphotransferases. The streptomycin phosphotransferases show some deviation from the

template, in that they do not have 3 hydrophobic amino acids following the asparagine. The template finds all of the protein kinases. With the exception of aph.6, all the sequences have the invariant asparagine, and all have the invariant pair of aspartic acids.

Most of the antibiotics are aminoglycosides, but viomycin is not, and the protein kinases have serine, threonine or tyrosine in peptide chains as their substrates. The conservation of the sequence suggests that this is a part of the protein that is involved in a common function; the obvious possibility is the binding of ATP and the phosphorylation of the substrate. The common pair of aspartic-acid residues could be involved in the binding through Mg^{2+} of the phosphate groups of ATP, and even in their activation. These and other possibilities are open to test by genetic manipulation of the sequences.

Other common sequence motifs have been reported. There may be many more and I am now engaged in trying to extract these automatically and exhaustively to see what sort of order this imposes on the database.

SYDNEY BRENNER

MRC Unit of Molecular Genetics,
MRC Centre, Hills Road,
Cambridge, CB2 2QH, UK

- Gray, G. & Fitch, W. *Molec. Biol. Evol.* **1**, 57–66 (1983).
- Herbert, C.J., Giles, I.G. & Akhtar, M. *FEBS. Lett.* **160**, 67–71 (1983).
- Beck, E., Ludwig, G., Auerswald, E.A., Reiss, B. & Schaller, H. *Gene* **19**, 327–336 (1982).
- Oka, A., Sugisaki, H. & Takanami, M. *J. molec. Biol.* **147**, 217–226 (1981).
- Thompson, C.J. & Gray, G. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5190–5194 (1983).
- Ferretti, J.J., Gilmore, K.S. & Courvalin P. *J. Bact.* **167**, 631–638 (1986).
- Bibb, M.J., Ward, J.M. & Cohen, S.N. *Molec. gen. Genet.* **199**, 26–36 (1985).
- Zalacain, M. et al. *Nucleic Acids Res.* **14**, 1565–1581 (1986).
- Gritz, L. & Davies, J. *Gene* **25**, 179–188 (1983).
- Kaster, K.R., Burgett, S.G., Rao, R.N. & Ingolia, T.D. *Nucleic Acids Res.* **11**, 6895–6911 (1983).
- Mazodier, P., Cossart, P., Giraud, E. & Gasser, F. *Nucleic Acids Res.* **13**, 195–205 (1985).
- Tohyama, H., Okami, Y. & Umezawa, H. *Nucleic Acids Res.* **15**, 1819–1833 (1987).
- Schwartz, D.E., Tizard, R. & Gilbert, W. *Cell* **32**, 853–861 (1983).
- Reimann, E.M. et al. *Biochemistry* **23**, 4185–4192 (1984).
- Van Bevern, C. *Nature* **289**, 258–262 (1981).

aph.1	V F S H G	D L G D S N I F V G -- D G -- K V S G F I	D L G R S G R A D K W Y D I A F C V R S I
aph.2	V F A H G	D Y C A P N L I I D -- G E -- K L S G F I	D C G R A G V A D R Y K D I S L A I R S L
aph.3	V V T H G	D A C L P N I M V D -- N G -- R F S G F I	D V G R L G V A D R Y Q D I A L A T R D I
aph.4	V V T H G	D F S L D N L I F D -- E G -- K L I G C I	D V G R V G I A D R Y Q D L A I L W N C L
aph.5	V V C H G	D L C P N N V L L D -- P G T C R V T G V I	D W G R L G V A D R H A D I A L A S R E L
aph.6	C L C H N D	D F S C N H L L L D -- G N N -- R L T G I I	D F G D S G I I D E Y C D F I Y L L E D M
vph.1	V V -- H G	D L G G E N V L W E T V D G V P R M S G V V	D W D E V G I G D P A F D L A A I G A S Y
hph.1	L V -- H A	D F G S N N V L T D L A A T -- R I T A V I	D W S E A M F G D S Q Y E F A N I F F W R
hph.1	F V -- H G	D L H G T N I F V D -- N G -- E V T G I V	D F T D V Y A G D S R Y S L V Q L H L N A
sph.1	P L -- H G	D L H H E N V D L F -- G D -- R G W L A I	D P H G L L -- G E R T F D Y A N I F T N P
sph.2	V L -- H W	D L H Y E N V D A A -- E A -- E P W L A I	D P E P L V -- A D P G F D L W F A L D T G
prk.1	Y V -- H R	D L R A A N I L V G -- E N -- L V C K V A D	D F G L A R L I -- -- E D N E Y * T A R Q G
prk.2	I V -- H R	D L K P E N I L L D -- D D -- M N I K L T	D F G F S C Q L D P G E K L R E V C G T P
prk.3	I L -- H L	D L K P A N I L I S -- E Q -- D V C K I S	D F G C S Q K L Q D L R G R Q A S P P N I

Enzymes: aph. (3'-aminoglycoside phosphotransferases): 1, *Streptococcus faecalis*¹; 2, *Bacillus circulans*²; 3, Tn5³; 4, Tn903⁴; 5, *Streptomyces fradiae*⁵; 6, 2'-aminoglycoside phosphotransferase⁶; vph. (vio-mycin phosphotransferase): 1, *Streptomyces vinaceus*⁷; hph. (hygromycin B phosphotransferase): 1, pJR225 (*E. coli*)⁸; 2, *Streptomyces hygroscopicus*^{9,10}; sph. (streptomycin phosphotransferase): 1, Tn5¹¹; 2, *Streptomyces griseus*¹²; prk. (protein kinases): 1, src (rous sarcoma virus)¹³; 2, gamma subunit of phosphorylase kinase (bovine)¹⁴; 3, v-mos mouse sarcoma virus¹⁵.

Electric charge of the neutrinos from SN1987A

SIR—The detection of nearly a dozen neutrinos, of 7 to 35 MeV in energy, coming presumably from the 1987 supernova in the Large Magellanic Cloud (LMC), 50 kpc away, all within some ten seconds¹, has been interpreted by several people as a proof that electron anti-neutrinos have a mass no larger than about 10eV.

Here we want to point out that, even if their mass is negligible, the bunching in time of the neutrinos indicates that their charge, q , is smaller than about 10^{-17} times the charge of the electron. If q were larger than this limit, the galactic magnetic field would lengthen their paths, and neutrinos of different energy could not arrive on the Earth within a few seconds of each other, even if emitted simultaneously by the supernova.

If B (tesla) is the uniform magnetic field normal to the path of the neutrino of charge q (in units of electronic charge e) and energy E (GeV), the radius of curvature of the trajectory (in metres) is

$$r = \frac{E}{0.3Bq}$$

The arc, $r\theta$, followed by the neutrino is longer than the straight path, s , by

$$\delta s = r\theta - 2r \sin \frac{\theta}{2} \approx s \frac{\theta^2}{24} \quad (\theta = \frac{s}{r} \ll 1)$$

so

$$\frac{\delta s}{s} \approx \frac{1}{24} \frac{(0.3Bs q)^2}{E^2}$$

The difference in length of the trajectories of two neutrinos of average energy E , differing in energy by ΔE , is then

$$\frac{\Delta s}{s} = \frac{\Delta t}{t} = \frac{1}{12} \frac{(0.3Bs q)^2}{E^2} \frac{\Delta E}{E}$$

where $\Delta t = s/c$ and Δt is the difference in the arrival time of the neutrinos on the Earth.

In conclusion,

$$\frac{q}{e} \leq \frac{\sqrt{12}}{0.3} \frac{E}{Bs} \left(\frac{\Delta t/t}{\Delta E/E} \right)^{1/2} \\ = 3.8 \times 10^{-12} \frac{(E/10 \text{ MeV})}{(s/10 \text{ kpc}) (B/1 \mu \text{G})} \left(\frac{\Delta t/t}{\Delta E/E} \right)^{1/2}$$

For the neutrinos from the LMC supernova, $E \approx 15$ MeV, $\Delta E/E \approx 1/2$ and $\Delta t \leq 5$ s.

Consider two cases: (1) an intergalactic field $B = 10^{-3} \mu\text{G}$ over the whole path, $s = 50$ kpc, $t = 5 \times 10^{12}$ s; (2) the galactic field, $B = 1 \mu\text{G}$, over a distance $s = 10$ kpc, $t = 1 \times 10^{12}$ s. In the first case, we find $q/e \leq 2 \times 10^{-15}$, and in the second, $q/e \leq 2 \times 10^{-17}$. (These examples consider homogeneous fields. A number, n , of randomly oriented inhomogeneities would increase our estimates by a factor of \sqrt{n} . In the