response within minutes) may initiate a signal cascade and mobilize calcium, resulting in auxin-specific responses.

We thank Drs W Tanner and H. Beevers for stimulating discussions, advice and help in preparing the manuscript. 3,5-D was kindly supplied by Dr M Zenk. This work was supported by the Deutsche Forschungsgemeinschaft.

Received 29 July; accepted 20 November 1987.

1. Hager, A., Menzel, H. & Krauss, A. *Planta* 100, 47-75 (1971).
2. Cleland, R. E. *A. Rev. Pl. Physiol.* 22, 197-222 (1971).
3. Felle, H., Brummer, B., Bertl, A. & Parish, R. W. *Proc. natn. Acad. Sci. U.S.A.* 83, 8992-8995 (1986).
4. Taiz, L. *A. Rev. Pl. Physiol.* 35, 585-657 (1984).
5. Hagen, G., Kleinschmidt, A. & Guilfoyle, T. *Planta* 162, 147-153 (1984).
6. Theologis, A., Huynh, T. V. & Davies, R. W. *J. molec. Biol.* 183, 53-68 (1985).
7. Löbler, M. & Klämbt, D. *J. biol. Chem.* 260, 9854-9859 (1985).
8. Berridge, M. J. *Biol. Chem. Hoppe-Seyler* 367, 447-456 (1986).
9. Nishizuka, Y. *Nature* 308, 693-698 (1984).
10. Micheli, B. *Nature* 319, 176-177 (1986).
11. Majerus, P. W. *et al. Science* 234, 1519-1526 (1986).
12. Ettlinger, C., Schindler, J. & Lehle, L. *Planta* 168, 101-105 (1986).
13. Murashige, T. & Skoog, F. *Physiologia Pl.* 15, 437-497 (1962).
14. Clarke, N. G. & Dawson, R. M. C. *Biochem. J.* 195, 301-306 (1981).
15. Irvine, R. F., Änggord, E. E., Letcher, A. J. & Downes, C. P. *Biochem. J.* 229, 505-511 (1985).
16. Turk, J., Wolf, B. A. & McDaniel, M. L. *Biochem. J.* 237, 259-263 (1986).
17. Hawkins, P. T., Stephens, L. & Downes, C. P. *Biochem. J.* 238, 507-516 (1986).
18. Berridge, M. J., Downes, C. P. & Hanley, M. R. *Biochem. J.* 206, 587-595 (1982).
19. Burgess, G. M., McKinney, J. S., Irvine, R. F. & Putney, J. W. Jr *Biochem. J.* 232, 237-243 (1985).
20. Bansal, V. S., Inhorn, R. C. & Majerus, P. W. *J. biol. Chem.* 262, 9444-9447 (1987).
21. Ackermann, K. E., Gish, B. G., Honchar, M. P. & Sherman, W. R. *Biochem. J.* 242, 517-524 (1987).
22. Irvine, R. F., Letcher, A. J., Heslop, J. P. & Berridge, M. J. *Nature* 320, 631-634 (1986).
23. Streb, H., Irvine, R. F., Berridge, M. J. & Schulz, I. *Nature* 306, 67-69 (1983).
24. Berridge, M. J. & Irvine, R. F. *Nature* 312, 315-321 (1984).
25. Hepler, P. K. & Wayne, R. O. *A. Rev. Pl. Physiol.* 36, 397-439 (1985).
26. Drobak, B. K. & Ferguson, I. B. *Biochem. biophys. Res. Commun.* 130, 1241-1246 (1985).
27. Rincon, M. & Boss, W. F. *Pl. Physiol.* 83, 395-398 (1987).
28. Schumaker, K. S. & Sze, H. J. *J. biol. Chem.* 262, 3944-3946 (1987).
29. Irvine, R. F., Letcher, A. J. & Dawson, R. M. C. *Biochem. J.* 192, 279-283 (1980).
30. Helsper, J. P. F. G., DeGroot, P. F. M., Linskens, H. F. & Jackson, J. F. *Phytochemistry* 25, 2053-2055 (1986).
31. Elliott, D. & Skinner, J. D. *Phytochemistry* 25, 39-44 (1986).
32. Sandelius, A. S. & Sommarin, M. *FEBS Lett.* 201, 282-286 (1986).
33. Heim, S., Bauleke, A., Wylegalla, C. & Wagner, K. G. *Pl. Sci.* 49, 159-165 (1987).
34. Boss, W. F. & Massel, M. O. *Biochem. biophys. Res. Commun.* 132, 1018-1023 (1985).
35. Heim, S. & Wagner, K. G. *Biochem. biophys. Res. Commun.* 134, 1175-1181 (1986).
36. Morse, M. J., Crain, R. C. & Satter, R. L. *Pl. Physiol.* 83, 640-644 (1987).
37. Morre, D. J., Gripshover, B., Monroe, A. & Morre, J. T. *J. biol. Chem.* 259, 15364-15368 (1984).

Functioning haemoglobin genes in non-nodulating plants

Didier Bogusz*†, Cyril A. Appleby*, Jörg Landsmann*‡, Elizabeth S. Dennis*, Michael J. Trinick* & W. James Peacock*

* Commonwealth Scientific and Industrial Research Organization, Division of Plant Industry, GPO Box 1600, Canberra, A.C.T. 2601, Australia
† ORSTOM Department MAA, 213 rue La Fayette, 75480 Paris, France
‡ Institute für Biochemie, Biologische Bundesanstalt, Messeweg 11/12, D-3300 Braunschweig, FRG

Haemoglobin has previously been recorded in plants only in the nitrogen-fixing nodules formed by symbiotic association between *Rhizobium* or *Frankia* and legume or non-legume hosts[1–4]. Structural similarities amongst these and animal haemoglobins at the protein and gene level suggested a common evolutionary origin[1,5–8]. This suggests that haemoglobin genes, inherited from an ancestor common to plants and animals, might be present in all plants. We report here the isolation of a haemoglobin gene from *Trema tomentosa*, a non-nodulating relative of *Parasponia* (Ulmaceae)[9]. The gene has three introns located at positions identical to those in the haemoglobin genes of nodulating plant species, strengthening the case for a common origin of all plant haemoglobin genes. The data argue strongly against horizontal haemoglobin gene transfer from animals to plants. The *Trema* gene has a tissue-specific pattern of transcription and translation, producing monomeric haemoglobin in *Trema* roots. We have also found that the *Parasponia* haemoglobin gene is transcribed in roots of non-nodulated plants. These results suggest that haemoglobin has a role in the respiratory metabolism of root cells of all plant species. We propose that its special role in nitrogen-fixing nodules has required adaptation of the haemoglobin-gene regulation pathway, to give high expression in the specialized environment of the nodule.

We have searched for haemoglobin genes in a number of plant species. Using a soybean haemoglobin complementary DNA probe, we were unable to detect (because of sequence divergence) the haemoglobin gene sequence even in *Parasponia*, a nodulating species known to produce haemoglobin[6]. We were able to isolate the *Parasponia* gene using a *Parasponia* cDNA clone identified with an oligonucleotide probe designed on the basis of amino-acid sequence data[6]. Because of the limitation of sequence divergence we first searched, using a *Parasponia* cDNA probe, for haemoglobin genes in non-nodulating plants in species of Ulmaceae, the family to which *Parasponia* belongs[9]. We chose *Trema tomentosa*, taxonomically close to *Parasponia*, and *Celtis australis*, located in another section of the family. *Trema* species do not have nodulating roots in their native habitats, and we have shown that *T. tomentosa* is not capable of forming symbiotic associations with any of 98 representative strains of *Rhizobium* and *Bradyrhizobium*. We were unable to detect any nodulation in another *Trema* species, *T. orientalis*, or in *Celtis*[9].

Southern hybridization showed haemoglobin-related sequences in both *Trema* and *Celtis*. The *Parasponia* haemoglobin cDNA segment hybridized to one band of *Trema* DNA following digestion with *Eco*RI or *Bgl*II. This suggests that there is a single haemoglobin-related segment in the genome of *Trema*, a situation comparable to *Parasponia* where there is only one haemoglobin gene. In *Celtis* the results indicate that there are either one or two genes present. Legume species have several haemoglobin genes and pseudogenes[10,11].

In *Trema*, we sought to determine whether the haemoglobin-related sequences correspond to a complete functional gene by isolating the appropriate DNA segment from the *T. tomentosa* genome (Fig. 1). Comparison of the *Trema* gene with the *Parasponia* gene[6] identified four exons and three introns in identical positions. The *Trema* and *Parasponia* exon sequences have 93% nucleotide similarity and there is ~80% similarity in the introns, non-translated leader and 3'-untranslated sequences of the genes.

A haem-polypeptide with amino-acid sequence deduced from the *Trema* gene sequence (Fig. 1) would have the properties expected of a functional haemoglobin. It has a chain length (161 residues) identical to *Parasponia* haemoglobin, and has the 30 amino-acid residues common to all plant haemoglobins whose structures are known[6–8,12,13]. These include the residues equivalent to proximal and distal His, Pro C2 and Phe CD1, which are invariant or highly conserved in all functional haemoglobins. Of the 11 amino acids not identical between the *Parasponia* and *Trema* polypeptides (Fig. 1), seven are not found in any other plant haemoglobin. These are Asp 6, Lys 34, Ile 58, Met 71, Arg 93, Asn 96 and Ser 154, with only Arg 93 having the potential to cause instability in *Trema* haemoglobin; compare with refs 8 and 14. We conclude that the *Trema* gene is likely to be functional, and that its product could participate in oxygen transport[1].

We analysed *Trema* tissue for RNA transcripts and for protein products of the gene (Fig. 2). Poly(A)$^+$ RNA extracted from *Trema* roots (lane 4) hybridized to *Parasponia* haemoglobin cDNA but no haemoglobin messenger RNA was detected in *Trema* leaf extracts (lane 5). The single band in root RNA indicates an mRNA of 0.7 kilobases (kb), identical in length to the haemoglobin mRNA in *Parasponia* nodules (lane 1).

```
                                                                                                                    100
TREMA CCCCTGGTTTTTCATACATATATGCATGCGTCATTAGCTA+++ACACCCCCTGAAGTTTTTTTTTCTTTTGGGGCTGCTAAGACC+TCATAAAGTTTACT
PARA         C                    C  A  G        AAAG T          A          T+++++          TTC   CC   G       T

                                                                                                                    200
TREMA AAAGATTAA+ACTTACTAAAAAGAAAACGAAAAAAAAAAAAAAAACCCAAGGAGATGGCTCTCCAGTACCCTGAAGAGTTACATCTTATGCTTATCCCCAT
PARA  C  A   TTAT                          T       ++++      A  T        C   A                    CACG   ++++++

                                                                                                                    300
TREMA TTTTT++CCTATATATACAGAGTGCCTTTACCAGATTTTCCAAACACACTCCAACATAATCCCATTGCCCAAATCAAA++TTCTCAGTTTTTAGTCCCCT
PARA        CTA '-'                  C                                  +                    A   AT       C

                                    HINCII                                                                          400
                          SerSerGluValAspLysValPheThrGluGluGlnGluAlaLeuValValLysSerTrpAlaValMetLysLysAsnS
TREMA CAACCCACAGAACCGATGAGCAGCTCAGAAGTTGACAAAGTTTTCACAGAAGAGCAGGAAGCTCTGGTGGTGAAATCATGGGCTGTAATGAAGAAGAACT
PARA              G C                  A                 G                    G
                                       Asn                                     Ala

                                                                                          ACCI                      500
      erAlaGluLeuGlyLeuLysPhePheLeuLys
TREMA CTGCTGAACTGGGTCTTAAATTCTTCCTCAAGTAAGTCAAGATTATATATGGTACACTTTTTATTTACTTTGCTTCTTTTGTAGACTAAGTTTTTGAATC
PARA                    C                    A          A                                A     C          A
                        Gln

                                                                                                                    600
                                                                                  IlePheGluIleAlaProSerAlaLysAsnLeuPheSerTy
TREMA AAAAGGTTTTTTTTCTTTCTTTTTTCTGAAAAAATTTGGTTGATTTGAAACTTTGCAGGATATTTGAGATTGCACCGTCTGCCAAGAACTTGTTCTCGTA
PARA     G   AC A   T   T C++++          A        +                                                        T

      rLeuLysAspSerProIleProLeuGluGlnAsnProLysLeuLysProHisAlaMetThrValPheValMet                                    700
TREMA TTTGAAGGACTCTCCGATTCCTTTGGAGCAGAACCCAAAGCTGAAGCCCCATGCTATGACTGTCTTCGTTATGGTAAAGCCAACTTTCGTTCTCCTATTC
PARA                  G                      C              C                                        T
                      Val                                   Thr

                                                                                                                    800
TREMA CCTTATCCTAATTTTTCCAAGAATCTAATGTTAATAAAATAGTATTTTGCCAATTTAAACAACCAAAAATTTAGACACAACTATATAAAATATTTTAATT
PARA                  A A +                                T                                        C

                                                                                                                    900
TREMA CTTGGGT+TTATGATACATTGATCTATAGTGATTGATCTATAATGATTCCAATTTCCCGTGTTACATATATGAGTTGTGCCAGCAACAATCGCATCACAG
PARA      T GT          C          A     +++++++++++++     C          G   T           T      G

                                                                                                                   1000
TREMA TCGCCTATTCCTATAAGCAGAATGGACGAC++TGACTGGTGAGACATATCACAGCAAAGCTCTACAATCTTTATGTTCCTCACTTGCTCTG+++TCTTTC
PARA      T       +++++++      A        TG      CT             A          A     T         T              TTC T

                                                                                                                   1100
                          ThrCysGluSerAlaValGlnLeuArgLysAlaGlyLysValThrValArgGluSerAsnLeuLysArgLeuGlyAlaI
TREMA TCTGGTACTG++CCTGGAAAGACGTGTGAATCTGCGGTTCAACTTCGGAAAGCCGGAAAAGTAACAGTGAGAGAATCAAACTTGAAAAGACTTGGGGCTA
PARA            TGT           A     G                                                G       A       G          A
                                                                                                 Lys     Asp        Ile

      leHisPheLysAsnGlyValValAsnGluHisPheGlu                                                                        1200
TREMA TCCACTTCAAAAATGGCGTAGTTAATGAACATTTTGAGGTACTACCCTGGCGACATAGTAGATATAATTCCATAAGTGTAATCCAAGCATTTGTTGTTTA
PARA              C                                        C  T              C              A
                Thr

                                                                                                                   1300
                                                                                                                    Val
TREMA GAGTCAAATTATTATTATTCTGTATGGTGGTTCTTGAATAATCGATCTTATTGTGGTATTTAAAAATTATATTATTCATGGGAAAAACGGTTTGTAGGTC
PARA                                                    A        CT          G              A

                                                                                                HINDIII BCLI      1400
      ThrArgPheAlaLeuLeuGluThrIleLysGluAlaValProGluMetTrpSerProGluMetLysAsnAlaTrpGlyGluAlaTyrAspGlnLeuValA
TREMA ACAAGGTTTGCACTTTTGGAGACCATAAAGGAAGCAGTACCAGAAATGTGGTCACCTGAGATGAAGAACGCATGGGGAGAAGCTTATGATCAGTTGGTTG
PARA                                                                                      T
                                                                                         Val

                                    SCAI                                                                            1500
      laAlaIleLysSerGluMetLysProSerSerThr
TREMA CTGCTATCAAGTCCGAAATGAAACCCTCCAGTACTTGAGAATTTTTATAGTTCTTGTAACATT+AGGTTTGAATAATGTGACAAAACTTATACTTA++++
PARA      C       T                                                  G    A TG                                  ATTA
                Phe

                                                                                                                   1600
TREMA CATTTGC+TGAGAGAAAGGTTATAACTTACATATTT++++GCCTTTTTATGTATCATAGTGTGACTC+ATCTCTCAACTTGTGTTGTTCATCTTGTTCAA
PARA  G     A       G     A     +  G     G GTATAA T GCA                        G A        C

TREMA AAGGAATTAGTCTTTCACTTTACATTTTGGG
```

Fig. 1 Alignment of the *T. tomentosa* (TREMA) and *P. andersonii* (PARA) haemoglobin genes. The amino-acid translation is given above and below the coding sequence. That of *Trema* is given in full. Only variant nucleotides and amino acids are given for *Parasponia*. Non-coding sequences are aligned to give maximum base-for-base similarity. The proposed TATA box is underlined (____). Crosses (+++) represent the sequences deleted or inserted.

Methods. A genomic library in λEMBL4 was made from a 4–7 kb *Bgl*II-fraction of *Trema* total DNA. About 2×10^5 recombinant phages were screened with ^{32}P-nick-translated *Parasponia* haemoglobin cDNA probe pL209 (ref. 6). Six genomic clones were isolated and one of them was subcloned successively in pUC18 and pBGS18 (ref. 19). The sequence was determined by subcloning fragments into M13 (mp18, 19) phages and sequencing by the dideoxy-chain termination method[20].

We also detected haemoglobin mRNA in non-nodulated *Parasponia* roots (Fig. 2, lane 2), the level of mRNA being about 1,000-fold less than that found in *Parasponia* nodule tissue (lane 1). No signal was found in *Parasponia* leaves (lane 3). The *Parasponia* haemoglobin gene thus has two differently-regulated patterns of expression, a high level of transcription in nodules and a low level of transcription in normal root tissue. We did not detect any haemoglobin mRNA in non-nodulated soybean roots with either soybean or *Parasponia* haemoglobin cDNA probes; compare with refs 15 and 16. If there is expression of haemoglobin genes in non-nodule tissue of soybean, either it is at a level too low to be detected or it is from a non-symbiotic gene sufficiently divergent in sequence that no cross hybridization occurs with the probes we used.

We fractionated *Trema* root extracts to look for haemoglobin protein. Western blot analysis (Fig. 3) showed that the haemoglobin mRNA is translated in *Trema* roots and that the native product is a monomer as it is in nodules of legumes and *Casuarina*. In *Parasponia* nodules the native haemoglobin is a dimer[2,17] and this is also true for the haemoglobin found in normal (non-nodulated) root tissue. Analysis of the root haemoglobin of *Parasponia* will determine if the monomer form is root-specific for that genus.

Our data suggest that all plants have haemoglobin genes, and

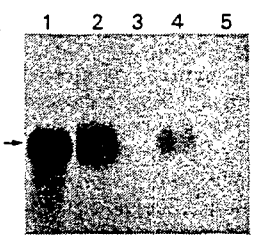

Fig. 2 Northern blot hydridization of *Parasponia* haemoglobin cDNA (pL209) (ref. 6) to total RNA isolated from *Parasponia* nodules (lane 1) and to poly(A)$^+$ RNA isolated from *Parasponia* roots (lane 2), *Parasponia* leaves (lane 3), *Trema* roots (lane 4) and *Trema* leaves (lane 5).

Methods. Uninoculated roots of *Parasponia andersonii* were obtained from plants grown aseptically on nutrient agar. Total RNA was extracted and poly(A)$^+$ purified according to standard procedures[21]. About 2 μg of total *Parasponia* nodule RNA and 5 μg of poly(A)$^+$ RNA was subjected to electrophoresis on a 1% formaldehyde agarose gel[21], blotted and hybridized. The probe was labelled with ^{32}P by an oligolabelling procedure[22]. The blots were hybridized at 42 °C as described[6] but with 10% dextran sulphate added. The filters were washed twice in a 2 × SSC, 0.1% sodium dodecyl sulphate (SDS) at room temperature for 15 min, then once in 0.1 × SSC, 0.1% SDS at room temperature for 30 min before autoradiography.

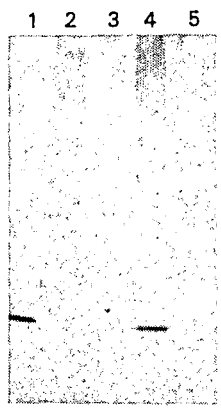

Fig. 3 Western blot identification of a native monomeric haemoglobin from *T. tomentosa* root tips, following non-denaturing molecular-exclusion chromatography on Sephacryl S200. Lane 1, pure *Parasponia* haemoglobin subunit, 0.1 μg; lane 2, the 'haemoglobin dimer' fractions $M_r \sim 39$K, ex S200; lane 3, the 'peroxidase' $M_r \sim 33$K fraction ex S200; lane 4, the 'haemoglobin monomer' fraction, $M_r \sim 18$K, ex S200; lane 5, the 'cytochrome *c*' $M_r \sim 13$K fraction ex S200. The arrow indicates subunit $M_r \sim 19$K. A significant interaction with the *Parasponia* haemoglobin antiserum occurred only in the 'haemoglobin monomer' fraction (lane 4). We consider that the faint bands seen in the other lanes result from adsorption of *Trema* monomeric haemoglobin to other proteins during the non-denaturing chromatography.

Methods. Frozen tissue (18 g) was extracted anaerobically and chromatographed on a calibrated column of Sephacryl S200 (Pharmacia, post 1980 batch) by a non-denaturing procedure[2,4]. The S200 effluent fractions expected to contain native dimeric haemoglobin ($M_r \sim 39$K)[2,17], peroxidase ($M_r \sim 33$K)[12], native monomeric haemoglobin ($M_r \sim 18$K)[4] and cytochrome *c* ($M_r \sim 13$K)[12] were separately pooled and concentrated to 1 ml. Residual polyvinyl pyrrolidone[4] was precipitated at 0.45 ammonium sulphate saturation, then putative haemoglobin and other proteins precipitated at 0.8 ammonium sulphate saturation. The redissolved protein pellets were water-washed and reconcentrated then proteins reprecipitated with 80% v/v acetone at −80 °C, and redissolved to ~100 μl in Western blot digestion buffer[4]. Samples were applied to SDS-polyacrylamide gel lanes before Western blot analysis[4] using IgG from rabbit anti-*Parasponia* haemoglobin serum, followed by protein A-peroxidase staining.

imply that haemoglobin has a function, presumably associated with oxygen transport, in cells of normal roots. In *Parasponia*, the product of a single gene sequence appears to have both root and nodule functions, but the level of gene expression is low in roots compared with nodules. We have not been able to detect haemoglobin mRNA in soybean roots but the possibility remains that there is a separate haemoglobin gene for the root function in legumes. Clearly if haemoglobin genes are present in all plants and encode proteins which function in root oxygen transport, then it is unnecessary to invoke horizontal gene transmission[5,18] to explain the sparsely-scattered appearance of symbiotic haemoglobins in nitrogen-fixing nodules[1-4]. Instead, the ubiquitous root haemoglobin gene may be the progenitor of the nodule genes. In this case we might expect that the nodule gene has evolved differently in each group of nodulating plants. In some species, for example *Parasponia*, the nodule and root haemoglobin functions may be provided by alternative regulation of the one gene, whereas in legumes, gene duplication may have been involved in the evolution of nodule-specific genes.

We thank Dr Don Spencer for help with Western blot analysis, Drs Jeff Ellis and Danny Llewellyn for helpful suggestions, Paul Hadobas for plants and Georgina Koci, Janice Norman and Judy Flanigan for technical assistance. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00296.

Note added in proof: Western blot analyses of *Parasponia* root haemoglobin showed it also to be a dimer.

Received 10 September; accepted 23 November 1987.

1. Appleby, C. A. *A. Rev. Pl. Physiol.* **35**, 443-478 (1984).
2. Appleby, C. A., Tjepkema, J. D. & Trinick, M. J. *Science* **220**, 951-953 (1983).
3. Tjepkema, J. D. *Can. J. Bot.* **61**, 2924-2929 (1983).
4. Fleming, A. I., Wittenberg, J. B., Wittenberg, B. A., Dudman, W. F. & Appleby, C. A. *Biochem. biophys. Acta* **911**, 209-220 (1987).
5. Hyldig-Nielsen, J. J. *et al. Nucleic Acids Res.* **10**, 689-701 (1982).
6. Landsmann, J. *et al. Nature* **324**, 166-168 (1986).
7. Fleming, A. I. *et al.* in *Proc. 8th Australian Nitrogen Fixation Conf.* (eds Wallace, W. & Smith, S. E.) 85-86 (Australian Institute of Agricultural Science, Parkville, 1986).
8. Dikerson, R. E. & Geis, J. *Hemoglobin: Structure, Function, Evolution and Pathology* (Benjamin/Cummings, California, 1983).
9. Akkermans, A. D. L., Abdulkadir, S. & Trinick, M. J. *Plant and Soil* **49**, 711-715 (1978).
10. Bojsen, K., Abildsten, D., Jensen, E. Ø., Paludan, K. & Marcker, K. A. *EMBO J.* **2**, 1165-1168 (1983).
11. Lee, J. S., Brown, G. G. & Verma, D. P. S. *Nucleic Acids Res.* **11**, 5541-5553 (1983).
12. *National Biological Research Foundation PIR Protein Sequence Database* release 12 (17 March, 1987).
13. Kortt, A. A., Strike, P. M., Bogusz, D. & Appleby, C. A. *Biochem. Int.* **15**, 509-516 (1987).
14. Fermi, G. & Perutz, M. F. *Haemoglobin and Myoglobin* (Clarendon, Oxford, 1981).
15. Marcker, A., Lund, M., Jensen, E. Ø. & Marcker, K. A. *EMBO J.* **3**, 1691-1695 (1984).
16. Jensen, J. S., Marcker, K. A., Otten, L. & Schell, J. *Nature* **321**, 669-674 (1986).
17. Wittenberg, J. B., Wittenberg, B. A., Gibson, Q. H., Trinick, M. J. & Appleby, C. A. *J. biol. Chem.* **261**, 13624-13631 (1986).
18. Runnegar, B. N. *J. molec. Evol.* **21**, 33-41 (1984).
19. Spratt, B. G., Hedge, P. J., Heesen, S. T., Edelman, A. & Broome-Smith, J. K. *Gene* **41**, 337-342 (1986).
20. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
22. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).

The mast cell binding site on human immunoglobulin E

Birgit Helm*, Philip Marsh*, Donata Vercelli†, Eduardo Padlan‡, Hannah Gould* & Raif Geha†

* Department of Biophysics, King's College, 26-29 Drury Lane, London WC2B 5RL, UK
† Division of Allergy and Immunology, Children's Hospital and Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA
‡ Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

Antibodies of the immunoglobulin E isotype sensitize mast cells and basophils for antigen-induced mediator release by binding through the Fc portion to a high-affinity receptor ($Fc_\varepsilon R1$, $K_a = 10^9\ M^{-1}$) on the cell surface[1,2] causing the clinical manifestations of type I hypersensitivity. As the amino acid sequence of the human epsilon chain is now known[3], attempts have been made to map the $Fc_\varepsilon R1$ binding site on IgE to a fragment smaller than Fc_ε using proteolytic cleavage products, none of which proved to be active[3]. Cleavage between the $C_\varepsilon 2$ and $C_\varepsilon 3$ domains released two inactive fragments, suggesting that the junction between these segments could be important in receptor binding[3]. This region is protected against protease digestion in the rat IgE complex with the receptor of rat basophilic leukaemia cells[4]. Here we report the mapping of the mast cell receptor binding site on human IgE to a sequence of 76 amino acids at the $C_\varepsilon 2/C_\varepsilon 3$ junction. Recombinant peptides containing this sequence inhibit passive sensitization of skin mast cells *in vivo* and sensitize mast cells to degranulation by anti-IgE *in vitro* almost as efficiently as a myeloma IgE. Fragments containing the separate domains are inactive. Additional sequences are required for rapid assembly of fragments into disulphide-linked dimers, suggesting that a single chain can form the active site. In a three-dimensional model of the human Fc_ε, the two identical segments are far apart. Each folds to generate a cleft between the $C_\varepsilon 2$ and $C_\varepsilon 3$ domains on the surface of the Fc_ε. The docking of IgE on to mast cells could take place within this cleft.

The bacterial expression product of the whole recombinant Fc_ε (rFc_ε) has been shown to be as active as a myeloma IgE in binding assays *in vivo* and *in vitro* and in assays of biological activity[5-9]. Here we extended this approach to the expression of smaller fragments of the ε-chain gene and assay their capacity to inhibit sensitization of mast cells by IgE antibodies *in vivo* and to sensitize basophils to degranulation by anti-IgE *in vitro*. The recombinant Fc_ε forms a disulphide-linked dimer on oxidation *in vitro*, resembling native Fc_ε (refs 5, 7), so we have examined the dimerization of the smaller peptides in relation to their activity. This information has been applied with a model of the three-dimensional structure of the human Fc_ε (ref. 10)