Pergamon

Tetrahedron Letters, Vol. 35, No. 5, pp. 765-768, 1994
Elsevier Science Ltd
Printed in Great Britain
0040-4039/94 $6.00+0.00

0040-4039(93)E0301-Y

A CONVENIENT SYNTHESIS OF *S*-CYANOETHYL-PROTECTED 4-THIOURIDINE AND ITS INCORPORATION INTO OLIGORIBONUCLEOTIDES

Chris J. Adams*, James B. Murray, John R. P. Arnold and Peter G. Stockley

Department of Genetics, University of Leeds, Leeds LS2 9JT, UK.

A reliable preparation of S^4-cyanoethyl-4-thiouridine and its incorporation into oligoribonucleotides is reported. Deprotection of oligoribonucleotides with DBU in acetonitrile followed by methanolic ammonia allows the use of standard N-benzoyl and N-isobutyryl protected phosphoramidites. Cleavage of hammerhead ribozymes using GCGCCGAAACACCGUG[^{4S}U]CUCGAGC as the modified substrate and GGCUCGACUGAUGAGGCGC as the ribozyme resulted in a halving of the cleavage rate when compared to the unmodified substrate, which is consistent with the proposal that the A^9-U^{17} base pair plays a key role in the active structure.

Synthetically, sulphur has been incorporated into the phosphate backbone[1,2], sugar[3] and bases[1] of oligonucleotides. Two major approaches for incorporating thio-bases into oligodeoxynucleotides, using the standard phosphoramidite approach, are currently in use. The first involves incorporation of either a *S*-cyanoethyl[4-6] or *S*-pivaloyloxymethyl[7,8] protected thiopyrimidine or thiopurine into the growing oligonucleotide, followed by deprotection. The other involves the incorporation of a suitably modified nucleotide which can be further altered after its incorporation into the oligonucleotide. Using this method a number of modified oligonucleotides can be prepared from the same synthesis[9]. 4-Thiopyrimidines and 6-thiopurines have been used for post-synthetic modification[10] and photo cross-linking[11,12].

The thione functionality also has different H-bonding properties to the carbonyl group and so can be used to probe protein-nucleic acid and oligonucleotide-oligonucleotide interactions.

Most of the above modifications have been on oligodeoxynucleotides. We report here the synthesis of S^4-cyanoethyl-4-thiouridine and to our knowledge the first reported incorporation of 4-thiouridine into a synthetic oligoribonucleotide, namely the 24-mer ribozyme substrate GCGCCGAAACACCGUG[^{4S}U]-CUCGAGC, which is selectively cleaved at C^{18}-U^{19} by a 19-nucleotide ribozyme[13] (Fig. 1). Its rate of cleavage is compared with that of the wild-type strand. In a preliminary experiment 4-thiouridine was also incorporated into the dimer [^{4S}U]G.

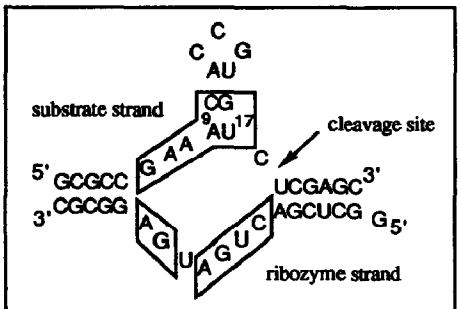
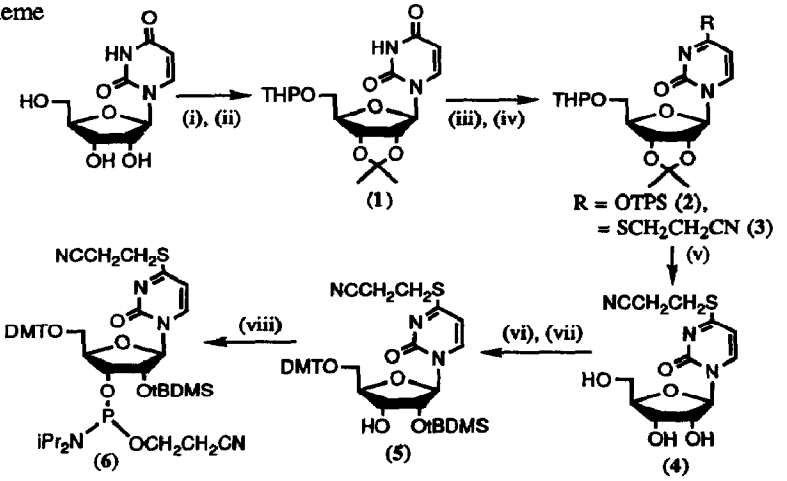


Figure 1 Hammerhead ribozyme. Nucleotides in boxes are essential for function

Coleman and Siedlecki [14] have reported the synthesis of S^4-cyanoethyl-protected 4-thiouridine and 2'-deoxy-4-thiouridine via reaction of the corresponding 4-tri-*O*-isopropylbenzenesulphonyl imidates with 3-mercaptopropionitrile (5 equiv) in aqueous ethanol (1.5 equiv K_2CO_3, 25°C, 3 h). Our attempts only afforded the corresponding hydrolysis products 2',3',5'-tri-*O*-trimethylsilyluridine or 2',3',5'-tri-*O*-*t*-butyldimethylsilyluridine. Our final approach is shown in the Scheme. 2',3'-*O*-Isopropylideneuridine[15] was tetrahydropyranylated (3,4-dihydro-2*H*-pyran, pyridinium *p*-toluenesulphonate, CH_2Cl_2, 1.5 h) to **1** in good yield (85%)[16]. We found that this protection strategy gave optimum results in the following steps. The reagents are cheap and the reactions proceed cleanly with good yields. A good yield (79 %) of the corresponding 4-*O*-tri-isopropylbenzenesulphonyl imidate (**2**) was achieved by the method of Gaffney *et al.*[17]. We could only achieve low yields using Bischofberger's[18] approach. Reaction of **2** with

$HSCH_2CH_2CN$[19] (1.2 equiv.) and *N*-methylpyrrolidine (1.5 equiv.) in CH_2Cl_2[20] produced the S^4-cyanoethyl derivative[21] in high yield (84 %). Introduction of the S^4-cyanoethyl group was also attempted using an approach reported by Connolly[4] who prepared S^4-cyanoethyl-protected 4-thiothymidine based on the method described by Miyasaka *et al.*[22] Unfortunately **3** was one of four products, being isolated in 15% yield. Deprotection of the acid-labile groups on the ribose was achieved with trifluoroacetic acid/water/CH_2Cl_2 (4.5:0.5:5). Deprotection was complete within 45 min at RT affording **4**[23] in high yield (86 %). Tritylation[24] and phosphitylation[25] to yield **6** were performed using standard procedures[26]. Unfortunately the selective 2'-silylation procedure developed by Ogilvie *et al.*[27] produced very low yields of **5**. Better yields (46%) were obtained using a standard silylation procedure [TBDMS-Cl (3 equiv), pyridine, 60 h][28]. The 3'-protected isomer was also obtained (16%) but was readily separated[27] using flash chromatography (diethyl ether/light petroleum 3:1).

Scheme



(i) 2,2-dimethoxypropane, *p*-toluenesulphonic acid, DMF (ii) 3,4-dihydro-2*H*-pyran, pyridinium *p*-toluenesulphonate, CH_2Cl_2 (iii) 2,4,6-tri-*O*-isopropylbenzenesulphonyl chloride, DMAP, iPr_2NEt, CH_2Cl_2 (iv) $HSCH_2CH_2CN$, *N*-methylpyrrolidine, CH_2Cl_2, (v) TFA, H_2O, CH_2Cl_2 (vi) 4,4'-dimethoxytrityl chloride, pyridine (vii) TBDMS-Cl, pyridine (viii) $iPr_2NP(Cl)OCH_2CH_2CN$, DMAP, iPr_2NEt, CH_2Cl_2

Incorporation of **6** into the dimer and 24-mer were carried out on an ABI 391PCR-Mate DNA synthesiser (Applied Biosystems). Phosphoramidites were from Chemgene Inc., Waltham, Mass., the bases being benzoyl-protected for C and A and isobutyryl-protected for G. CPG-benzoylG and -benzoylC were from Millipore. The *S*-cyanoethyl ether and *O*-cyanoethyl phosphate esters were deprotected using 0.3 M DBU in acetonitrile for 1 h at room temperature[4]. Cleavage from the support and deprotection of the exocyclic amine protecting groups was achieved in a sealed container with freshly prepared methanolic ammonia (methanol at 0°C purged with NH_3 for 30 min) at 30°C for 24 h. After evaporation the residue was 2'-*O*-TBDMS deprotected using either $NEt_3.3HF$ for 16 h (dimer)[29] or 1 M TBAF in THF (24-mer) at 30°C for 30 h. Poor solubility of the 24-mer in $NEt_3.3HF$ precluded its use. Based on trityl assays the overall yield of the 24-mer was 67.5% (average stepwise yield 98.3%) with incorporation of ^{4S}U lower at 90%. Both componds were purified by HPLC[30] using a linear gradient of acetonitrile from 2-10% over 40 min. Buffer A was 0.1 M ammonium acetate pH 6.5, buffer B 50% acetonitrile and 0.05 M ammonium acetate, pH 6.5. Two major products were isolated, the required modified 24-mer along with the unmodified analogue. The slightly longer retention time of the modified 24-mer is consistent with the presence of a more hydrophobic sulphur atom. No separation was observed using PAGE. This is in agreement with Favre *et al.*[8] who required a combination of agarose affinity chromatography and affinity electrophoresis to separate a 2'-deoxy-4-thiouridine-containing 14-mer from non-thiolated 14-mer. The

final yield after HPLC purification was 416 μg or 7.5% based on the final trityl assay[31]. Approximately 200 μg of unmodified 24-mer was also isolated. The dimer was prepared with a 92% coupling step. The final purified yield was approximately 400 μg (65 %) based on trityl assay. It is possible that better yields would be obtained using the faster base-deprotecting PAC (Pharmacia)[32] or FOD (ABI)[33] phosphoramidites.

The UV spectra of the 24-mer and dimer (Fig. 2) contain peaks at 342 and 338 nm, respectively, which is characteristic of the thione chromophore. Composition of the 24-mer was confirmed by HPLC (Fig. 3) after base composition analysis[34]. When recorded at 332 nm one peak was observed whose retention time corresponded to an authentic sample of 4-thiouridine. Similar retention times for ^{4S}U and G required that separate HPLC traces be recorded at 260 and 342 nm rather than simply recording one and changing the observation wavelength when ^{4S}U eluted. Figure 3 shows a superimposition of the two traces.

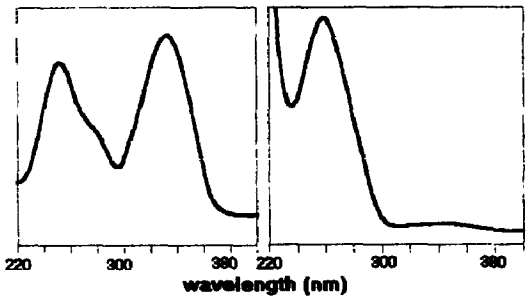


Figure 2. UV spectra of GCGCCGAAACACCGUG-[^{4S}U]CUCGAGC and [^{4S}U]G

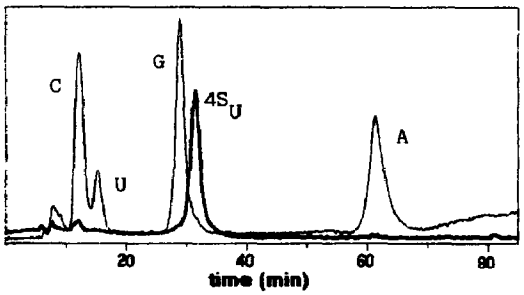


Figure 3. HPLC analysis after base composition analysis of GCG-CCGAAACACCGUG[^{4S}U]CUCGAGC. Traces recorded at 260 nm for C, U, G and A. Separate trace recorded at 332 nm for ^{4S}U (injection sample six-times larger for this trace).

A number of modified oligoribonucleotides have been prepared in order to study the mechanism of hammerhead ribozyme cleavage[35]. Replacement of exocyclic oxygen with sulphur introduces a bulkier and weaker H-bonding atom but the same ring structure is retained. Single turnover reactions, carried out in duplicate, were performed at 37°C in a volume of 40 μL 50 mM Tris-HCl, pH 7.3, with 5'-^{32}P-labelled substrate at a concentration of 0.2 μM and a ribozyme concentration of 1 μM. Cleavage was intiated by the addition of $MgCl_2$ to a final concentration of 10 mM. Aliquots of 4 μL were removed at 1 min intervals and quenched with 5 μL of 8 M urea and 50 mM EDTA. The samples were analysed by PAGE on 20% denaturing gels and the resulting autoradiographs subjected to laser scanning densitometry. The half-life for cleavage of the modified strand was 450 ± 20 s compared with the unmodified strand which was 240 ± 10 s. This represents a doubling of the half-life and is consistent with the H-bond between the A^9-U^{17} base pair being important in the active structure. Slim and Gait[36] showed that the ribozyme had lost most of its activity when the exocyclic amine was removed from A^9.

In conclusion, S^4-cyanoethyl-4-thiouridine was prepared via a reliable and cheap route. Its incorporation into a dimer and 24-mer oligoribonucleotide was performed in reasonable yields especially given the use of non-fast-deprotecting phosphoramidites. The modified 24-mer was cleaved at half the rate of the wild-type strand.

This work was supported by grants from the SERC and the Wellcome Trust. We would also like to thank Dr. J Fisher for NMR spectra and Prof. B. Connolly, University of Newcastle, for useful advice.

References and Notes:

1. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 1925.
2. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543.
3. Bellon, L.; Barascut, J.-L.; Maury, G.; Divita, G.; Goody, R.; Imbach, J.-L. *Nucl. Acids Res.* **1993**, *21*, 1587.
4. Nikiforov, T. T.; Connolly, B. A. *Tetrahedron Lett.* **1992**, *33*, 2379.
5. Connolly[4] has used other protecting groups but the S-cyanoethyl has proved to be the best.

6. Waters, T. R.; Connolly, B. A. *Nucleosides and Nucleotides* **1992**, *11*, 985.
7. Clivio, P.; Fourrey, J.-L.; Gasche, J.; Favre, A. *J. Chem. Soc. Perkin Trans. I* **1992**, 2383.
8. Clivio, P.; Fourrey, J.-L.; Gasche, J.; Audic, A.; Favre, A.; Perrin, C.; Woisard, A.; *Tetrahedron Lett.* **1992**, *33*, 65
9. Xu, Y.-Z.; Swann, P. F.; Zheng, Q. *Tetrahedron* **1992**, *48*, 1729.
10. Coleman, R. S.; Siedlecki, J. M.; *J. Am. Chem. Soc.* **1992**, *114*, 9229.
11. Milton, J.; Connolly, B. A.; Nikiforov, T. T.; Cosstick, R. *J. Chem. Soc. Chem. Commun.* **1993**, 779.
12. Woisard, A.; Favre, A.; Clivio, P.; Fourrey, J.-L. *J. Am. Chem. Soc.* **1992**, *114*, 10072.
13. Uhlenbeck, O. C. *Nature* **1987**, *328*, 596.
14. Coleman, R. S.; Siedlecki, J. M.; *Tetrahedron Lett.* **1991**, *32*, 3033.
15. Chládek, S.; Smrt, J. *Coll. Czech. Chem. Comm.* **1963**, *28*, 1301.
16 ^{13}C NMR: δ ($CDCl_3$) 19.50, 25.09, 30.33 (3 × t, THP CH_2), 25.31, 27.18 (2 × q, isopropylidene CH_3), 62.56, 62.64 (t, 5'-C), 67.12, 67.40 (t, THP OCH_2), 80.60, 80.23 (d, 2'-C), 85.04, 85.28 (d, 3'-C), 85.53, 85.61 (d, 4'-C), 92.50, 92.61 (d, 1'-C), 98.83, 99.44 (d, THP OCH), 101.75, 102.03 (d, 5-C), 113.94, 114.13 (s, isopropylidene $\underline{C}Me_2$), 140.64, 141.28 (d, 6-C), 150.13 (s, 4-C), 163.51 (s, 2-C).
17. Gaffney, B. L.; Marky, L. A.; Jones, R. A. *Tetrahedron* **1984**, *40*, 3.
18. Bischofberger, N. *Tetrahedron Lett.* **1987**, *28*, 2821.
19. Bauer, L.; Welsh, T. L. *J. Org. Chem.* **1961**, *26*, 1443.
20. Rao, T. S.; Jayaraman, K.; Durland, R. H.; Revankar, G. R.*Tetrahedron Lett.* **1992**, *33*, 7651.
21. ^{13}C NMR: δ ($CDCl_3$) 18.20 (t, $\underline{C}H_2CN$), 19.47, 19.74 & 25.07, 25.17 & 30.28, 30.44 (3 × t, THP CH_2), 25.28, 27.18 (2 × q, isopropylidene CH_3), 25.31 (t, SCH_2), 62.65, 62.97 (t, OCH_2), 66.96, 67.39 (t, THP OCH_2), 80.40, 80.56 (d, 2'-C), 85.87, 86.12 (d, 3'-C), 86.79, 86.93 (d, 4'-C), 94.33, 94.53 (d, 1'-C), 98.71, 99.79 (d, THP OCH), 102.55, 102.88 (d, 5-C), 113.69, 113.85 (s, isopropylidene $\underline{C}(CH_3)_2$), 118.06 (s, C≡N), 141.18, 141.89 (d, 6-C), 153.26, 153.32 (s, 2-C), 175.40, 175.55 (s, 4-C).
22. Matsuda, A.; Obi, K.; Miyasaka, T. *Chem. Pharm. Bull.* **1985**, *33*, 2575.
23. ^{1}H NMR: δ (d^6-dmso) 2.96 (2H, t, J 4.8 Hz, CH_2CN), 3.38 (2H, J 4.8 Hz, SCH_2), 3.62 (1H, d, J 9.3 Hz, 5'-H), 3.76 (1H, d, J 9.2 Hz, 5'-H), 3.93, 4.00 (3H, 2 × s, 2',3',4'-H), 5.75 (1H, s, 1'-H), 6.51 (1H, d, J 5.0 Hz, 5-H), 8.34 (1H, d, J 5.0 Hz, 6-H). ^{13}C NMR: δ (d^6-dmso) 21.50 (t, $\underline{C}H_2CN$), 28.54 (t, SCH_2), 63.70 (t, 5'-C), 72.53 (d, 2'-C), 78.65 (d, 3'-C), 88.24 (d, 4'-C), 94.28 (d, 1'-C), 106.76 (d, 5-C), 123.09 (s, C≡N), 146.00 (d, 6-C), 156.86 (s, 4-C), 178.76 (s, 2-C).
24. ^{1}H NMR: δ ($CDCl_3$) 2.87 (2H, t, J 5.0 Hz, CH_2CN), 3.39 (2H, t, J 5.1 Hz, SH_2), 3.43-3.50 (2H, m, 5'-H), 3.79 (6H, s, OCH_3), 4.33-4.41 (3H, m, 2',3',4'-H), 5.84 (1H, d, J 2.1 Hz, 1'-H), 5.97 (1H, d, J 5.3 Hz, 6-H), 6.82 (4H, dd J 6.6, 1.2 Hz, arom. H), 7.15-7.34 (9H, m, arom. H), 8.12 (1H, dd, J 5.3, 1.1 Hz, 6-H). ^{13}C NMR: δ ($CDCl_3$) 18.20 (t, $\underline{C}H_2CN$), 25.36 (t, SCH_2), 55.22 (q, 2 × OCH_3), 62.16 (t, 5'-C), 70.78 (d, 2'-C), 76.43 (d, 3'-C), 85.14 (d, 4'-C), 86.94 (s, DMT $O\underline{C}Ar_3$ Ar = C_6H_5(4-OMe-C_6H_4)$_2$), 92.84 (d, 1'-C), 103.88 (d, 5-C), 113.19 (d, arom. CH), 118.03 (s, C≡N), 127.05, 127.92, 127.94, 129.05, 129.96 (d, arom. CH), 134.97, 135.15 (s, arom. ipso C), 140.70 (d, 6-C), 144.09 (s, arom. ipso C), 154.85 (s, C(2)=O), 158.57 (s, arom. ipso C), 175.86 (s, 4-C).
25. ^{1}H NMR: δ ($CDCl_3$) 0.17, 0.29 (6H, 2 × s, $SiCH_3$), 0.91 (9H, s, Si{$C(CH_3)_3$}), 1.11, 1.15 (12H, 2 × d, NCH($C\underline{H}_3$)$_2$), 2.42 (2H, t, J 6.3 Hz, CH_2CN), 2.90, (2H, q, J 5.7 Hz, SCH_2), 3.32-3.74 (m, 5'-H, $N\underline{H}(CH_3)_2$, $OC\underline{H}_2CH_2$), 3.81 (6H, s, OCH_3), 4.32-4.42 (3H, m, 2',3',4'-H), 5.71 (s, 1'-H), 5.72 (d, J 7.2 Hz, 5-H), 6.85 (4H, d, J 8.4 Hz, PhH), 7.27-7.49 (9H, m, PhH), 8.42 (1H, d, J 7.2 Hz, 6-H).
26. Connolly, B. A., Oligonucleotides and Nucleotides A Practical Approach, Ed. Eckstein, F., Oxford University Press, New York, 1991, p 161, 175.
27. Hakimelahi, G. H.; Proba, Z. A.; Ogilvie, K. K. *Can. J. Chem.* **1982**, *60*, 1106.
28. ^{1}H NMR: δ ($CDCl_3$) 0.22, 0.35 (6H, 2 × s, $SiCH_3$), 0.94 (9H, s, Si{$C(CH_3)_3$}), 2.89 (2H, td J 4.9, 1.8 Hz, CH_2CN), 3.33, 3.44 (2H, 2 × m, SCH_2), 3.57 (2H, m, 5'-H), 3.80 (6H, s, OCH_3), 4.08 (1H, dt, J 6.4, 1.4 Hz, 4'-H), 4.30 (1H, d, J 3.5 Hz, 3'-H), 4.41 (1H, m, 2'-H), 5.74 (1H, d, J 5.3 Hz, 5-H), 5.79 (1H, s, 1'-H), 6.85 (4H, d, J 6.6 Hz, arom. H), 7.23-7.41 (9H, m, arom. H), 8.36 (1H, d, J 5.3 Hz, 6-H). ^{13}C NMR: δ ($CDCl_3$) -5.43, -4.33 (2 × q, $SiCH_3$), 18.06 (s, $Si\underline{C}(CH_3)_3$), 18.33 (t, $\underline{C}H_2CN$), 25.28 (t, SCH_2), 25.83 (q, $SiC(\underline{C}H_3)_3$), 55.24 (q, 2 × OCH_3), 60.86 (t, 5'-C), 68.64 (d, 2'-C), 76.43 (d, 3'-C), 83.04 (d, 4'-C), 87.00 (s, DMT $O\underline{C}Ar_3$), 90.91 (d, 1'-C), 103.29 (d, 5-C), 113.24 (d, DMT arom.C), 118.10 (s, C≡N), 127.09, 127.96, 128.12, 130.06, 130.11 (d, DMT arom. C), 135.00, 135.29 (s, DMT. ipso arom. C), 141.04 (d, 6-C), 144.20 (s, DMT ipso arom. C), 153.49 (s, 2-C), 158.61 (s, DMT ipso arom. C(OMe)), 175.30 (s, 4-C).
29. Gasparutto, D.; Livache, T.; Bazin, H.; Duplaa, A.-M.; Guy, A.; Khorlin, A.; Molko, D.; Roget, A.; Teoule, R. *Nucl. Acids Res.* **1992**, *20*, 5159.
30. Murray, J. B.; Collier, A. K.; Arnold, J. R. P. *Anal. Biochem.* in press.
31. The unmodified 24-mer is obtained in 12% yield.
32. Wu, T.; Ogilvie, K. K.; Pon, R. T. *Nucl. Acids Res.* **1989**, *17*, 3501.
33. Andrus, A.; Vu, H.; McCollum, C.; Jacobson, K.; Theisen, P.; Vinayak, R.; Spiess, E. *Nucl. Acids Res.* **1992**, *20*, 1265.
34. Eadie, J. S.; McBride, L. J.; Efcavitch, J. W.; Hoff, L. B.; Cathcart, R. *Anal. Biochem.* **1987**, *165*, 442.
35. Seela, F.; Mersmann, K.; Grasby, J. A.; Gait, M. J. *Helv. Chim. Acta* **1993**, *76*, 1809. and references therein.
36. Slim, G.; Gait, M. J. *Biochem. Biophys. Res. Comm.* **1992**, *183*, 605.

(*Received in UK* 22 *October* 1993; *revised* 22 *November* 1993; *accepted* 26 *November* 1993)