Physicochemical characterization of the cytoplasmic domain of the epidermal growth factor receptor and evidence for conformational changes associated with its activation by ammonium sulphate

Mary GREGORIOU,*†, Pam F. JONES,*§, John F. TIMMS,* Jenny J. YANG,† Sheena E. RADFORD† and Anthony R. REES*||

*Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, The Rex Richards Building, South Parks Road, Oxford, OX1 3QU, U.K. and

†Oxford Centre for Molecular Sciences, University of Oxford, New Chemistry Laboratory, Oxford, OX1 3QT, U.K.

The physicochemical properties of the purified cytoplasmic domain of the epidermal growth factor (EGF) receptor, its self-phosphorylation and peptide phosphorylation activities, and its activation by ammonium sulphate have been studied. Highly efficient purification procedures for the isolation of the recombinant cytoplasmic domain (Met⁶⁴⁴–Ala¹¹⁸⁶) of the EGF receptor, expressed in the baculovirus/insect cell system, are described. Physicochemical characterization of the protein included investigation of its isoelectric and hydrodynamic properties, stability, oligomeric status, and secondary structure using far-u.v. circular dichroism. The recombinant protein was not recognized by anti-phosphotyrosine antibodies, unless first self-phosphorylated *in vitro*. Tryptic phosphopeptide maps of self-phosphorylated recombinant cytoplasmic domain and the EGF-stimulated A431-membrane receptor were very similar, suggesting that the recombinant had similar self-phosphorylation capacity and specificity. The preparations were characterized by high specific activity towards peptide tyrosine phosphorylation. Although the cytoplasmic domain was isolated as a homogeneously monomeric protein, storage at 4 °C led to slow, spontaneous aggregation

with reduction in specific activity. Both high activity and monomeric state were maintained by storage below 0 °C. The dependence of the initial rate of self-phosphorylation on protein concentration was consistent with cross-phosphorylation but not with the known oligomerization-induced activation of holoreceptor. The peptide phosphorylation activity was stimulated by Mn²⁺, Mg²⁺ and (NH₄)₂SO₄ at high concentrations. The substrate specificity of (NH₄)₂SO₄ activation was studied using synthetic peptides. Self-phosphorylation was inhibited by (NH₄)₂SO₄ in the range 0–0.25 M but activated at 1.0–1.5 M, possibly as a result of ionic and hydrophobic protein interactions respectively. Phosphopeptide maps of cytoplasmic domain phosphorylated in the presence of high (NH₄)₂SO₄ showed that the protein was more extensively phosphorylated than in the absence of salt, or than the native receptor. Far-u.v. circular-dichroism spectra of the cytoplasmic domain changed dramatically at 1 M (NH₄)₂SO₄, raising the possibility that (NH₄)₂SO₄ activates the kinase catalytic domain by inducing conformational changes.

INTRODUCTION

Regulation of the protein tyrosine kinase activity of growth factor receptors is of primary importance in the control of cell growth and differentiation. The biochemical responses of cells to stimulation with growth factor involve complex interactions between growth-factor receptors and key components of mitogenic signalling pathways regulated by protein phosphorylation and dephosphorylation events (Yarden and Ullrich, 1988; Bohmer et al., 1993; Fry et al., 1993). The epidermal growth factor (EGF) receptor is a single-chain transmembrane glycoprotein which comprises an extracellular EGF-binding domain, a short transmembrane peptide and a cytoplasmic protein tyrosine kinase domain, the activity of which is essential for the function of the molecule (Ullrich et al., 1984; Chen et al., 1987; Honegger et al., 1987). The cytoplasmic domain consists of a juxtamembrane region, a catalytic subdomain which is closely homologous to the *v-erbB* oncogene product (Downward et al., 1984b) and a C-terminal region which carries the major self-phosphorylation

sites (Downward et al., 1984a; Hsuan et al., 1989; Margolis et al., 1989). EGF binds to the extracellular domain and stimulates kinase activity (Erneux et al., 1983; Downward et al., 1985) towards self-phosphorylation and phosphorylation of cellular substrates.

The mechanism of activation of self-phosphorylation of the EGF receptor is unclear. Experimental evidence supports an allosteric activation model in which receptor monomers upon EGF binding associate to form oligomers with higher affinity for EGF and stimulated kinase activity (Boni and Pilch, 1987; Yarden and Schlessinger, 1987; Canals, 1992). It is not known whether activating conformational changes occur in the cytoplasmic domain as a result of oligomerization. Although much evidence supports a dimerization-induced activation of native receptor, a recombinant, soluble cytoplasmic domain preparation has recently been shown to consist of monomers and dimers of similar activities, and aggregates of lower activity (Cadena et al., 1994). This study also demonstrated that self-phosphorylation of monomers and dimers was associated with conformational

Abbreviations used: DTT, dithiothreitol; EGF, epidermal growth factor; i.e.f., isoelectric focussing; m.o.i., multiplicity of infection; p.f.u., plaque-forming units; rHER, rabbit anti-EGF receptor antibody; SH2, src homology region 2; TFA, trifluoroacetic acid.

† To whom correspondence should be addressed.

§ Present address: Regeneron Pharmaceuticals Inc., 777 Old Saw Mill River Road, Tarrytown, NY 10591-6707, U.S.A.

|| Present address: Department of Biochemistry, University of Bath, Bath BA2 7AY, U.K.

extension of the cytoplasmic domain, which may be a significant step in receptor signalling through phosphotyrosine-src homology region 2 (SH2) domain interactions. Self-phosphorylation of the EGF receptor does not appear to increase its intrinsic activity (Downward et al., 1985; Cadena et al., 1994), unlike tyrosine self-phosphorylation of the insulin receptor, which is required for kinase activation.

Studies of kinase activation showed that Mn^{2+} and $(NH_4)_2SO_4$, at high concentrations, stimulate the kinase activity of the receptor under conditions where the protein is largely monomeric, suggesting that oligomerization is not a pre-requisite for the stimulation of kinase activity *in vitro* (Koland and Cerione, 1988; Northwood and Davis, 1988). A recent report has proposed that the mechanism of activation by $(NH_4)_2SO_4$ involves aggregation of the cytoplasmic domain, based on the finding that correlation existed between precipitation by structurally diverse polymers or precipitating agents [such as poly(ethylene glycol), polylysine, $(NH_4)_2SO_4$, etc.] and kinase activation by these agents (Mohammadi et al., 1993). Alternatively, conformational changes or direct interactions of sulphate with substrates at the active site have been suggested as possible mechanisms of kinase activation by $(NH_4)_2SO_4$ or Na_2SO_4 based on a kinetic study (Hubler et al., 1992). Although the physiological significance of activated monomeric receptors is not known, it is possible that alternative activating agents may act by similar mechanisms, or that activated monomers and oligomers share a common structural basis for activation, and thus warrant further study.

Ultimately, to understand the detailed structural basis of EGF receptor signalling it is necessary to determine the three-dimensional structure of the protein. The expression and purification system described here yields sufficient quantities of active cytoplasmic domain for crystallization studies. The aims of this study were twofold; to better define the physical properties and the behaviour of the cytoplasmic domain in preparation for structural studies, such as X-ray diffraction; and to further our understanding of its function with respect to oligomerization, and the mechanism of kinase activation by $(NH_4)_2SO_4$.

EXPERIMENTAL

Materials

Spodoptera frugiperda insect cell lines Sf9 and Sf21, *Autographa californica* nuclear polyhedrosis virus clones AcNPV and AcUW1-lacZ, and transfer vectors pAc373 (Summers and Smith, 1987) and pAcYM1 (Matsuura et al., 1987) were obtained from the Institute of Virology and Environmental Medicine (IVEM), Oxford. Cell-culture reagents were from Gibco and JRH Scientific. Oligonucleotides and peptides DREYVHPF and RRLIEDAEYAARG were from the Oxford Centre for Molecular Sciences (OCMS), DRVYVHPF from Bachem, (Bubendorf, Switzerland), EGVYVHPV from Sigma, and peptides RREELQDDYE and RREELQDDYEDD were gifts from Tim Page (Dyson Perrins Laboratory, Oxford). Antibody PY72 was a gift from Dr. J. Hayles (Imperial Cancer Research Fund, London). $(NH_4)_2SO_4$ low in heavy metals was from Fisons U.K. Ltd., and ACS-grade $(NH_4)_2SO_4$ was from Sigma.

Cell culture

Human A431, and insect Sf9 and Sf21 cells were grown as previously described (Gregoriou and Rees, 1984; Summers and Smith, 1987). Insect cells were maintained at $(0.1-2.0) \times 10^6$ cells/ml in TC-100 medium/10% (v/v) foetal calf serum, in spinner bottles, at 28 °C. High-titre [10^8 plaque-forming units (p.f.u.)/ml] virus stocks were prepared from Sf21 spinner cultures

infected at a multiplicity of infection (m.o.i.) of < 1 for 4-5 days. For protein purification, cells were grown in batches of several 300 ml cultures. Cells were spun at 800 rev./min (70 g) in a Sorvall GSA rotor, gently resuspended in virus containing medium at an m.o.i. of ≥ 10 , grown for 48 h in 3% foetal calf serum, harvested and stored at -70 °C until required. A431 cell membranes were prepared as described (Thom et al., 1977).

Cloning and construction of expression clones of the cytoplasmic domain

A cDNA library was constructed using poly(A)⁺ mRNA isolated from A431 cells under optimized conditions (Jones, 1987). cDNA was synthesized from poly(A)⁺ RNA and cloned into λ gt10 (NBL). Positive clones were identified by hybridization to a 500 bp *Bam*HI fragment from v-erbB cDNA or to oligonucleotide probes hybridizing to the second and fourth *Eco*RI fragment of the complete cDNA. *Eco*RI fragments (1.8, 0.77 and 1.3 kbp) from corresponding λ DNA clones were subcloned into pUC18 and sequenced (in M13K19 using the dideoxynucleotide sequencing method) to confirm their identity with the corresponding regions of the A431 EGF receptor (Ullrich et al., 1984). The *Eco*RI fragments were ligated in a series of steps. Site-directed mutagenesis (Carter et al., 1985) was used to modify *Bam*HI sites in the cytoplasmic coding region without altering the amino acid sequence, and new *Bam*HI cloning sites were introduced 5' to the ATG encoding Met⁶⁴⁴ and 3' to the *Sca*I site in the 3' untranslated region. The resulting cDNA (1.8 kb) encoding the cytoplasmic domain was sequenced to ensure that no spurious mutations had been introduced, and was cloned into the *Bam*HI sites of transfer vectors pAc373 and pAcYM1 at the *Bam*HI site downstream of the polyhedrin promoter. Sense clones pAc373EGFR/cy and pAcYM1EGFR/cy were selected in *E. coli* strain MC1061. Plasmid DNA was purified on CsCl gradients, and AcNPV and AcUW1-lacZ viral DNA was prepared as described (Overton et al., 1987). The AcUW1-lacZ genome contained the β -galactosidase coding sequence under control of the P10 promoter (Weyer et al., 1990) to facilitate viral titration using 5-bromo-4-chloro-3-indolyl- α -D-glucopyranoside (X-Gal).

Recombinant baculovirus clones were generated using standard methods (Overton et al., 1987; Summers and Smith, 1987). After co-transfection with vectors and wild-type virus clones, recombinants were isolated by limiting dilution and DNA hybridization (Fung et al., 1988). Recombinants were identified using a 1.3 kb *Pst*I cDNA probe labelled by the random primer method (Amersham system RPN.1600). After three cycles of cloning pure recombinants were isolated from a 1×10^{-8} dilution.

Detection of cytoplasmic domain expression in infected cells

Sf9 monolayers (2.5×10^5 cells) were infected with recombinant virus at m.o.i. ≥ 10 and incubated for 24, 48 and 72 h at 28 °C. Proteins were labelled with [³⁵S]methionine (14 μ Ci in 200 μ l of serum- and methionine-free TC100) for 3 h and analysed by SDS/PAGE and autoradiography. Alternatively, cytoplasmic domain was detected by Coomassie Blue staining of SDS gels.

Protein analysis by SDS/PAGE and isoelectric focussing (I.e.f.)

Proteins were analysed in polyacrylamide (10% w/v) gels as described (Laemmli, 1970), and visualized by Coomassie Blue, silver staining, immunoblotting or autoradiography using Fuji RX Medical Film at -70 °C. Immunoblots were incubated with rabbit anti-(EGF receptor) antibody (rHER) (1:2000, v/v),

followed by alkaline phosphatase-IgG, and were developed with 0.005 % (w/v) nitroblue tetrazolium and 0.0025 % (w/v) bromochloroindolyl phosphate. Protein samples (1 mg/ml, 25 μ l) were analysed by i.e.f., under non-denaturing conditions, in 4 % (w/v) polyacrylamide gels containing either pH 3–10 or pH 4–6 ampholyte base, in a BioRad Biophoresis Cell.

Assay of kinase activity

Enzyme activity was measured by following the incorporation of 32 P from [γ - 32 P]ATP into either [Val⁵]angiotensin II (DRVYVHPF) or the src peptide RRLIEDAEYAARG using the phosphocellulose paper binding assay (Pike, 1987). In assays which included peptide EGVYVHPV an alternative method was used (Braun et al., 1983). Kinase was assayed routinely using [γ - 32 P]ATP (17–20 μ M, 0.3–3.0 μ Ci) and 1 mM peptide in 30 mM Hepes/NaOH buffer, pH 7.4, containing 10 mM MnCl₂, at 4 or 22 °C (where stated). Phosphorylation of peptide was linear with time for at least 30 min at 4 °C and 12 min at 22 °C. Self-phosphorylation assays were similar except that peptide was excluded; reactions were stopped with 20 μ l of Laemmli sample buffer and analysed by SDS/PAGE and autoradiography. Sodium orthovanadate (130 μ M) was included in assays of crude kinase. $V_{\max}$ and apparent K_m (K_m^{app}) values were determined from initial rates measured at 22–23 °C using the direct linear plot, or averaged rates (from triplicate assays) were fitted to the Michaelis–Menten equation, using the programme LEONORA (written by A. J. Cornish-Bowden).

Purification of the cytoplasmic domain

Sf9 cells infected with recombinant baculovirus for 48 h were spun gently, resuspended in lysis buffer (35 ml of lysis buffer/1 of culture) and stored frozen at –70 °C. [Lysis buffer: 10 mM triethanolamine/HCl, pH 7.6, 5 mM NaEDTA, 5 mM dithiothreitol (DTT), 1 % (v/v) glycerol and 0.05 % (v/v) Triton X-100.] Infected cell cultures (total volume 8 l) were homogenized in the presence of the protease inhibitors leupeptin (12.5 μ g/ml), aprotinin (6.25 μ g/ml), E-64 (2.5 μ g/ml), pepstatin (1 μ g/ml) and 3,4-dichloroisocoumarin (2.5 μ g/ml). All purification steps were carried out at 4 °C except for the f.p.l.c. steps (20–22 °C). NaCl was added to a concentration of 100 mM, and a 100 000 g fraction of soluble protein was prepared. Cell debris was washed once with NaCl (100 mM, 125 ml) in buffer A (25 mM Hepes/NaOH, pH 7.4, 5 mM NaEDTA, 2.5 mM DTT, 5 % glycerol), and the wash was pooled with the 100 000 g fraction to a total volume of 400 ml. The cytoplasmic domain was purified by using two procedures concurrently.

Procedure 1

Soluble extract (100 ml, corresponding to 2 l of culture) was fractionated by (NH₄)₂SO₄ precipitation (0.105 g/ml extract, ~ 0.8 M). Precipitated cytoplasmic domain was dissolved in buffer A containing 100 mM NaCl, and purified on an SP Sepharose column (100 ml). The column was washed with 150 mM NaCl (250 ml), a linear gradient of NaCl (150–500 mM in 300 ml), and finally 500 mM NaCl. Cytoplasmic domain self-phosphorylation activity was associated with a peak of A_{280} at the end of the gradient [Figure 2a(i)]. The SP Sepharose fraction (95 ml) was concentrated to 7.7 ml at 0.25 M NaCl, and was loaded onto a Sephacryl S100 Hilo column (300 ml, Pharmacia Ltd) equilibrated in buffer A containing 150 mM NaCl. Three fractions containing the highest activity [Figure 2a(ii)] were combined. After threefold dilution with buffer A, the protein was finally

purified by f.p.l.c. on a Mono Q HR10/10 column (Pharmacia Ltd) equilibrated in buffer A containing 50 mM NaCl and 1 mM NaEDTA. The column was washed with equilibration buffer (12 ml) and developed with a linear gradient of NaCl (50–420 mM in a total volume of 60 ml). The cytoplasmic domain was recovered at ~ 200 mM NaCl in two closely eluting peaks of protein, I and II, which had kinase activity [Figure 2a(iii)]. Peaks I and II were pooled separately, where possible, and stored in 30 % (v/v) glycerol at –20 °C.

Procedure 2

Soluble extract (300 ml) from 6 l of culture was loaded directly onto an SP Sepharose column (200 ml) in buffer A containing 50 mM NaCl. The column was washed with 600 ml of buffer A containing 150 mM NaCl, and developed with a gradient of NaCl (250–500 mM in a total volume of 500 ml), followed by 500 mM NaCl (300 ml). Cytoplasmic-domain kinase activity was recovered in a broad peak of A_{280} near the end of the gradient, in 218 ml [Figure 2b(i)]. This fraction was diluted twofold with buffer A, and concentrated by ultrafiltration to 64 ml. A small amount of insoluble material formed and was removed by spinning for 30 min at 100 000 g. The concentrated protein was further chromatographed on an f.p.l.c. Mono S HR5/5 column (Pharmacia Ltd) in successive 20–25 ml portions (results not shown). The cytoplasmic domain bound to the column at 200 mM NaCl, and eluted isocratically at 300 mM NaCl as a broad peak. The Mono S fraction was concentrated to 8.3 ml (19 mg of protein) and chromatographed on Sephacryl S100, as for Procedure 1 [Figure 2b(ii)]. The cytoplasmic domain (13.4 ml) was diluted with 2 vol. of buffer A and purified further by f.p.l.c. on Mono Q, essentially as in Procedure 1, except that when the gradient reached 200 mM NaCl the protein was eluted isocratically. A broad peak (20 ml) of cytoplasmic domain was obtained [Figure 2b(iii)], which was resolved by analytical Mono Q chromatography into two closely eluting peaks, I and II, as in Procedure 1 (Figure 4a). DEAE-Sepharose could be used instead of Mono Q without compromising protein purity, yield, or resolution in peaks I and II.

Protein quantification

The Bradford assay (Bradford, 1976) was used to determine total protein in impure fractions. Purified protein was quantified on the basis of its A_{280} , using an experimentally derived absorption coefficient ($A_{280}^{1\%,1\text{cm}} = 9.19$). To determine $A_{280}^{1\%,1\text{cm}}$, the absorbance at 280 nm was measured from u.v. spectra of purified protein in buffer A, and samples (5–10 μ l of protein) subjected to quantitative amino acid analysis. The resulting averaged value for the amount of protein from four amino acid analyses was related to the corresponding $A_{280}^{1\%,1\text{cm}}$ using a calculated M_r of 61 308.

Sucrose-density-gradient sedimentation

Sedimentation analysis was carried out using sucrose density gradients [4.5 ml, 5–20 % (w/v) sucrose in 50 mM Hepes/NaOH buffer (pH 7.4)/0.5 mM NaEDTA/2.5 mM DTT/3 mM NaN₃]. Protein samples (100 μ l) in the same buffer were loaded onto the gradients and spun at 40 000 rev./min (150 000 g) in a Beckman SW55 Ti rotor at 20 °C for 15 h. Gradients were fractionated into 100 μ l portions on ice, and 10 μ l samples were assayed for self-phosphorylation activity. BSA (50 μ g) and carbonic anhydrase (200 μ g) protein standards were analysed simultaneously.

Calculation of physical constants

Sedimentation coefficients, s , were determined from the distance of migration of the protein from the meniscus of sucrose gradients, according to the method of Martin and Ames (Martin and Ames, 1961), using known values for the sedimentation coefficients of protein standards BSA and carbonic anhydrase (Smith, 1970). The Stokes (molecular) radius, a , of the cytoplasmic domain was determined by gel filtration on Sephacryl S100 calibrated with standard proteins, as described (Ackers, 1964). The molecular weight of the cytoplasmic domain was calculated from a and s , using the equation

$$M_r = (6\pi\eta N a s) / (1 - \bar{v}r)$$

where N is Avogadro's number, $\bar{v}$ is the partial specific volume (0.73), r is the density of water at 20 °C, and η is the viscosity of water at 20 °C as described (Siegel and Monty, 1966). The frictional coefficient f/f_0 was calculated from the equation

$$f/f_0 = a / [(3\bar{v}M_r)/(4\pi N)]^{1/3}$$

as described (Siegel and Monty, 1966).

Preparation of tryptic phosphopeptide maps

Self-phosphorylated cytoplasmic domain was extracted from an SDS gel, and digested with trypsin as described (Beemon and Hunter, 1978; Cooper et al., 1983). [32 P]-labelled proteins were extracted in 50 mM NH_4HCO_3 , pH 7.3–7.6, precipitated with trichloroacetic acid (20%, w/v) in the presence of RNAase carrier protein, oxidized with performic acid and extensively digested with trypsin (Promega sequencing grade; trypsin was added twice, each time at a trypsin/total protein ratio of 20:100, w/w) at 37 °C for 24 h. Samples were freeze-dried, dissolved in 100 μ l of trifluoroacetic acid (TFA) (0.1%, v/v) and soluble peptides were analysed by h.p.l.c., essentially as described (Hsuan et al., 1989). Samples (100 μ l) were resolved on a C18 Vydak column (5 μ m pore size, 25 cm) equilibrated in 0.1% TFA. The column was washed, and developed with a linear gradient of 0.96% min^{-1} acetonitrile in 0.1% TFA, at 1 ml/min, and fractions (0.3 ml) were counted in a scintillation counter.

Circular dichroism (c.d.)

Cytoplasmic domain was transferred to 50 mM Tris/ H_2SO_4 buffer (pH 7.5) by gel filtration on an f.p.l.c. Superose 12 column, stored on ice, and used within 8–12 h. The protein was diluted with equilibration buffer and equilibrated at 10 °C for 5 min in a 1 mm quartz cell. Far-u.v. c.d. spectra were recorded over the wavelength range 195–260 nm, at 0.5 nm intervals, using a Jasco J720 spectropolarimeter. For each of the protein samples eight spectra were averaged, and difference spectra were reported in molar ellipticity, $[\theta]_M$ (degrees $\cdot \text{cm}^2 \cdot \text{dmol}^{-1} \cdot \text{residue}^{-1}$).

The c.d. spectra of the cytoplasmic domain were fitted by reference to a set of spectra of 25 proteins, using the programme CCA (convex constraint analysis), and secondary-structure contents were determined with LINCOMB (Perczel et al., 1992). Forty iterations and four pure components were used to generate reference sets with minimum errors. Four pure components were generated as a reference set with conformational weights; fitting with five components did not result in significant reduction in the error.

The spectrum of component I was interpreted as helix, having two negative maxima at 208 and 222 nm; that of component II resembled β -sheet or β -turn, having a positive maximum at 200 nm and a negative maximum at 228 nm; that of component III was characteristic of random conformation, having one negative maximum at 198 nm; and that of component IV

displayed a large positive ellipticity at 230 nm and a small peak at 208 nm, and was assigned to β -turns and non-secondary-structure chiral contributions (e.g. aromatic side chains) (Perczel et al., 1992).

RESULTS

Expression of the cytoplasmic domain in *S. frugiperda* cells

The cDNA encoding the cytoplasmic domain of the EGF receptor was constructed from A431 clones hybridizing to known EGF-receptor DNA sequences (Ullrich et al., 1984), and cloned into vectors pAc373 and pAcYM1 at the *Bam*HI site, 3' to the polyhedrin promoter, with ATG encoding Met⁶⁴⁴ serving as a translation initiation codon. Recombinant baculovirus clones from four different assortments of vectors and virus clones (cytoplasmic domain clones pAc373EGFR/cy and pAcYM1EGFR/cy, and AcNPV or AcUW1-lacZ baculovirus clones) were cloned by limiting dilution and DNA hybridization and their yields of recombinant protein were compared. pAcYM1-derived recombinants produced at least a fivefold higher level of recombinant protein, but no significant variation within the pAcYM1EGFR/cy or pAc373EGFR/cy clone subsets was found. pAcYM1/AcUW1-lacZ-derived AcEGFR/cy clones were therefore selected for use.

The cytoplasmic domain had an apparent M_r of 62000–65000 on SDS polyacrylamide gels, where it was detected as a discrete band in extracts of recombinant but not wild-type virus-infected cells (Figure 1). This agrees well with the theoretical value of 61300. The cytoplasmic domain was also recognized by antiserum rHER (lanes 1 and 2) and by its self-phosphorylation activity (results not shown). Expression of the cytoplasmic domain was maximal at 48 h post-infection, as shown by [35 S]methionine bio-synthetic labelling studies (Figure 1). The level of expression was estimated from the intensity of the band corresponding to the cytoplasmic domain relative to total protein in the total cell extract,

Figure 1 Expression of cytoplasmic domain

(a) Sf9 cells (n, non-infected; w, wild-type AcNPV-infected; r, recombinant Ac373EGFR/cy-infected) were labelled for 3 h with [35 S]methionine 24, 48 and 72 h post-infection and total protein (12 μ g) was analysed by SDS/PAGE and autoradiography. The position of the cytoplasmic domain is indicated by arrowheads. (b, lanes 1 and 2) Soluble protein (32 μ g) from AcNPV-infected (lane 1) and Ac373EGFR/cy-infected (lane 2) cells harvested at 48 h and analysed by immunoblotting using rHER antiserum. (b, lanes 3 and 4) Comparison of expression levels of the cytoplasmic domain obtained at 48 h post-infection with recombinant virus clones Ac373EGFR/cy (lane 4) and AcYM1EGFR/cy (lane 6). Cytoplasmic domain was detected on Coomassie-Blue-stained SDS gels, by reference to non-expressing control cells infected with AcNPV (lane 3) and AcUW1 (lane 5).

Figure 2 Purification of the cytoplasmic domain

(a) Purification procedure 1. The $(\text{NH}_4)_2\text{SO}_4$ fraction of the cytoplasmic domain was chromatographed on (i) SP Sepharose, followed by (ii) Sephacryl S100, and (iii) Mono Q columns. (b) Purification procedure 2. The 100 000 *g* fraction of the cytoplasmic domain was chromatographed on (i) SP Sepharose (followed by Mono S (not shown)), (ii) Sephacryl S100, and (iii) Mono Q columns. Fractions indicated with crossed symbol were analysed for self-phosphorylation activity; the autoradiographs obtained are shown (insets).

analysed by SDS/PAGE and Coomassie Blue staining. In small cultures (2–5 ml) the recombinant was ~ 5 % of total cell protein, whereas in 300 ml cultures the expression level was lower [1–2 %; see Figures 1b (lane 6) and 3a (lane 1)]. On both scales, expression level was critically dependent on infecting cells at high m.o.i. (≥ 10). An accurate determination of the level of expression from measurements of kinase activity was difficult because in crude extracts kinase activity appeared depressed (either through inhibition, or dilution of radiolabel with endogenous ATP). The level of expression also depended critically on the time of harvest post-infection, which was optimal at 48 h (at m.o.i. ≥ 10), whereas later harvest times (52–72 h) caused considerable degradation. A catalytically active, 42 kDa fragment was obtained from a 56 h post-infection harvest in one of our initial preparations (results not shown). Phospho amino acid analysis of the ^{32}P -self-phosphorylated kinase domain showed that phosphotyrosine was the major phospho amino acid, as expected (results not shown).

Purification of the cytoplasmic domain

The cytoplasmic domain was produced in Sf9 cells after synchronous infection with the baculovirus clone AcEGFR/cy. The efficiency of extraction of the recombinant protein was determined qualitatively, as quantification of activity in crude extracts was difficult (see above). Figure 2 shows that most of the expressed protein was extracted in the soluble fraction, as judged by its intensity in whole-cell and 100 000 *g*-soluble extracts. However, qualitative estimations of soluble and insoluble frac-

tions in Western blots indicated that extraction was ~ 60–80 % efficient. Analysis of small-scale preparations indicated the presence of low-mobility components (presumably phosphorylated protein) in the non-extractable fraction.

Highly purified kinase domain was prepared by two similar procedures. Procedure 1 consisted of fractionation with $(\text{NH}_4)_2\text{SO}_4$, cation-exchange chromatography on SP Sepharose, gel filtration on Sephacryl S100, and anion-exchange chromatography on Mono Q (Figure 2a). Procedure 2 consisted of cation-exchange chromatography first on SP Sepharose followed by f.p.l.c. on Mono S, gel-filtration and Mono Q chromatography (Figure 2b). The Mono Q step partially resolved the protein into two peaks, I and II. Both procedures are described in the Experimental section, and the results are illustrated on silver-stained and rHER-immunostained SDS gels (Figure 3a). Densitometric analysis at 1 μg protein loading of Coomassie-Blue-stained gels showed that the cytoplasmic domain was at least 95 % pure. Similar yields of protein were obtained on a 2–8 l scale (1.5 mg/l of cells, as shown by amino acid analysis), of which 50–65 % was isolated as peak I. Procedures 1 and 2 required 3 and 4 days respectively; minimization of time may be important for the isolation of active, monomeric protein (see below). Preparations of ~ 90 % purity were obtained in an earlier two-step procedure, consisting of $(\text{NH}_4)_2\text{SO}_4$ precipitation and Mono Q HR10/10 chromatography, on a 1–2 l culture scale, which was completed in 1 day (Figure 3c). $(\text{NH}_4)_2\text{SO}_4$ (0.8 M) precipitation and cation-exchange chromatography on SP (or S) Sepharose in Procedures 1 and 2 provided highly efficient initial purification steps (Figure 3a). Cation-exchange chromatography was particularly useful in

Figure 3 Purity and immunoreactivity of the preparations of the cytoplasmic domain

(a) Silver-stained gel of samples from successive purification stages. Lane 1, homogenate (20 μ g); lane 2, 100 000 g -soluble extract (20 μ g). Purification procedure 1: lane 3, $(\text{NH}_4)_2\text{SO}_4$ supernatant (6 μ g); lane 4, $(\text{NH}_4)_2\text{SO}_4$ precipitate (6 μ g); lane 5, SP Sepharose (1.2 μ g); lane 6, S100 (1 μ g); lane 7, Mono Q peak I (1 μ g); lane 8, Mono Q peak II (1 μ g). Purification procedure 2: lane 9, SP Sepharose (8 μ g); lane 10, Mono S (1.3 μ g); lane 11, S100 (1 μ g); lane 12, Mono Q peak I and II (1 μ g). (b) Immunoblot corresponding to the above series of purification fractions, developed with rHER antiserum. (c) (i) Coomassie-Blue-stained gel of 10 Mono Q fractions across peaks I and II, from a two-stage purification procedure; the first 10 samples were not self-phosphorylated, whereas the second set were analysed after self-phosphorylation *in vitro*. (ii) Immunoblot corresponding to the above series of purification fractions, developed with rHER antiserum.

removing degradation products arising from proteolysis of the kinase in the juxtamembrane region, which was observed in some preparations.

Physicochemical characterization

Biosynthetic studies showed that the cytoplasmic domain was synthesized as a single type of polypeptide, which on SDS/PAGE behaved as a singlet protein band (Figure 1). Ion-exchange-purified preparations of the cytoplasmic domain were partially resolved into two close protein peaks, I and II, the resolution of which was increased by isocratic elution at 170 mM NaCl. This

Figure 4 Charge heterogeneity of monomeric cytoplasmic domain

(a) Resolution of Mono Q peaks I and II by isocratic elution at 150 mM followed by 170 mM NaCl (dotted line). (A minor peak of kinase activity eluting with 150 mM salt was not investigated.) (b) I.e.f. analysis of the cytoplasmic domain prepared according to Procedure 1. Unresolved peaks I and II, and peak II were analysed in the pH range 4–6, and were stained with silver. The positions of markers of known pI are indicated. (c) Sucrose-density-gradient sedimentation analysis of the cytoplasmic domain. Self-phosphorylation activity profiles are shown for Mono Q peak I (0.68 mg/ml, $\square$) and peak II (0.48 mg/ml, $\diamond$) prepared by Procedure 1, and for Mono Q peak I (0.09 mg/ml, Δ) and unresolved peaks I + II (0.03 mg/ml, $\circ$) prepared by Procedure 2. CA, carbonic anhydrase.

also revealed further apparent heterogeneity under both peaks I and II, and indicated the presence of a third, minor component, eluting at 150 mM NaCl (Figure 4a). SDS/PAGE analysis revealed further protein heterogeneity within peak II, which migrated as a doublet recognized by rHER antibody (Figure 3b, lane 8). Examination of the anti-phosphotyrosine reactivity of the preparations with monoclonal antibody PY72 revealed a minor trace of antibody reactivity (which might arise from non-specific staining) above the bands of fast mobility; this was equally distributed across peaks I and II [Figure 3c(ii)]. Therefore, the band of slow mobility in the doublet of peak II was not a phosphorylated species. *In vitro* self-phosphorylation resulted in equally strong antiphosphotyrosine reactivity in peaks I and II, suggesting that they contained similar amounts of phosphorylatable tyrosine. Furthermore, *in vitro* self-

Table 1 Kinetic parameters of the cytoplasmic domain

Cytoplasmic domain was prepared by Procedure 1 (Mono Q Peak I), and by Procedure 2 (Mono Q Peaks I and II), and parameters were determined at 22–23 °C. Values were either calculated from averaged rates measured in triplicates and fitted to the Michaelis–Menten equation, using the programme LEONORA, or determined using the direct linear plot. Data represent means $\pm$ S.D. ($n = 3$ –6). *The concentration of ATP was maintained at 64 mM as the concentration of peptide was varied. †The concentration of [Val⁵]angiotensin II peptide was maintained at 3.4 mM as the concentration of ATP was varied. —, Not determined.

Methods	$V_{\max}$ (pmol $\cdot$ min ⁻¹ $\cdot$ pmol ⁻¹)			K_m^{app} (mM)		
	Angiotensin II*	Src peptide*	ATP†	Angiotensin II*	Src peptide*	ATP†
Procedure 1, LEONORA	2.43 $\pm$ 0.33	0.96 $\pm$ 0.05	—	2.90 $\pm$ 0.88	0.94 $\pm$ 0.12	—
Procedure 1, linear plot	2.35 $\pm$ 0.35	1.75 $\pm$ 0.35	0.95 $\pm$ 0.02	2.18 $\pm$ 0.38	0.44 $\pm$ 0.03	0.012 $\pm$ 0.008
Procedure 2, LEONORA	2.76 $\pm$ 0.18	2.35 $\pm$ 0.13	—	1.02 $\pm$ 0.19	0.40 $\pm$ 0.06	—

Figure 5 Dependence of the initial rates of self-phosphorylation on protein concentration

Initial rates were measured over the first 10 s of the reactions on ice, as described; data represent means $\pm$ S.E.M. ($n = 3$). Experiments were completed within 24 h of repeated gel filtration of monomeric protein, to ensure that aggregates were not present. (a) Non-linear dependence of rate on protein concentration. (b) The stoichiometry of self-phosphorylation over the initial 10 s of the reaction indicates that $\sim 2.6\%$ of the maximal phosphorylation capacity was observed. (c) Logarithmic plot of initial rates against protein concentration (see text for details).

phosphorylation showed that all fractions across the two peaks were phosphorylated to a similar extent [as shown in Figure 2a(iii)]. Therefore, peaks I and II do not represent different tyrosine phosphorylation states of the cytoplasmic domain.

Analysis of the isoelectric properties of protein under peaks I and II by i.e.f. revealed charge heterogeneity. In the pH range 3.5–9.3 the protein focussed as a broad band at pH 5–6. Fractions

from the ascending side of peak I focussed mostly as a narrow band, though additional more acidic minor bands were detected (results not shown). Analysis of the unresolved Mono Q peak and of peak II in the pH range 4–6, and visualization with silver staining, showed that both samples focussed in the pH range 5.1–5.4 and consisted of three major and several minor components (Figure 4b).

Electrospray m.s. revealed that protein from the ascending side of peak I (prepared by Procedure 1) consisted of a major species of M_r 61 381.56 $\pm$ 15.84, higher than the calculated average mass of 61 308 by 73. Secondary peaks corresponded to species with additional masses of 96 and 134, the first of which could be attributed to strongly bound SO_4^{2-} . The reasons for these discrepancies in mass were not investigated. All detectable phosphorylating activity was associated with a major protein peak during gel filtration (Figure 2b), which eluted close to monomeric BSA, suggesting that the preparation might be monomeric.

Sucrose-density-gradient sedimentation analysis showed that the cytoplasmic domain migrated as a single symmetrical peak of self-phosphorylation activity (Figure 4c). Protein from Mono Q peaks I and II, as well as the unresolved Mono Q peak, prepared by either procedure and tested over a 10-fold concentration range (0.09–0.7 mg/ml), migrated the same distance from the meniscus of the gradients. This indicated that heterogeneity did not arise from protein oligomerization. The Stokes (molecular) radius, a , the sedimentation coefficient, $s_{20,w}$, the M_r and the frictional ratio, f/f_0 , were determined experimentally or calculated as described in the methods. Values for $s_{20,w}$ were in the range 3.73–4.23 S [mean $\pm$ S.E.M. ($n = 4$) 3.98 $\pm$ 0.25 S]. The Stokes radius was determined from the elution position of the cytoplasmic domain by gel filtration on Sephacryl S100, which was calibrated with standards, as described (Ackers, 1964); the mean value for protein prepared by Procedures 1 and 2 was 3.74 $\pm$ 0.08 nm. Using these values for $s_{20,w}$ and a , the calculated M_r of the cytoplasmic domain was 62 559, indicating that it was monomeric. The calculated frictional ratio f/f_0 was 1.54, larger than the values for globular proteins [≤ 1.3 (Smith, 1970)], suggesting that the cytoplasmic domain is elongated.

The purified EGF-receptor cytoplasmic domain formed large aggregates on storage at 4 °C but not at -20 or -70 °C. Small amounts of aggregates were detected as early as 48 h after purification, and formed with protein from either peak, irrespective of the type of purification procedure used. Aggregated protein was detected by gel filtration on f.p.l.c. Superose 12 or Superose 6 columns (Pharmacia Ltd), where it eluted in the void volume, suggesting an apparent M_r of $> 1 \times 10^6$. No intermediate-size oligomers were observed, and the process of aggregation could not be reversed by dilution of the protein.

Figure 6 Stimulation of kinase activity by $(\text{NH}_4)_2\text{SO}_4$.

(a) $(\text{NH}_4)_2\text{SO}_4$ stimulation of the $[\text{Val}^5]$ angiotensin II phosphorylation activity of intact cytoplasmic domain (○) and of an active 42 kDa catalytic fragment (□) obtained by calpain digestion. (b) Stimulation of phosphorylation of $[\text{Val}^5]$ angiotensin II analogues by $(\text{NH}_4)_2\text{SO}_4$. Rates were measured at 0.5 mM peptide, 20 μM ATP and at 4 °C, using the assay described by Braun et al. (1983). The rate values presented are averages of duplicated assays. ○, DRVYVHPF; □, EGVYVHPV; and Δ, DREYVHPF. Similar results were obtained in three other experiments. (c) Inhibitory and activating effects of $(\text{NH}_4)_2\text{SO}_4$ (○) and NaCl (□) on the rate of self-phosphorylation of the cytoplasmic domain.

Characterization of peptide- and self-phosphorylation activities

The activity of the monomeric cytoplasmic domain towards peptide phosphorylation was characterized using the widely used $[\text{Val}^5]$ angiotensin II (DRVYVHPF) and the src peptide analogue RRLIEDAEYAARG. V_{max} and K_m^{app} values for the two peptides and for ATP were determined. These kinetic parameters did not vary appreciably with the method of preparation or with different protein isoforms (Table 1), and are similar to those reported previously (Hsu et al., 1990; McGlynn et al., 1992; Cadena et al., 1994). Activity towards $[\text{Val}^5]$ angiotensin II and src peptide was markedly decreased upon aggregation (see below). The cytoplasmic domain also catalysed phosphorylation of the $[\text{Val}^5]$ angiotensin II homologues DREYVHPF and EGVYVHPV, and peptides RREELQDDYE and RREELQDDYEDD [erythrocyte band-3 phosphorylation site (Kemp and Pearson, 1991)].

In order to characterize the function of the aggregated state of the cytoplasmic domain, the self-phosphorylation activities of the isolated monomeric and aggregated forms were compared. At protein concentrations of 12 $\mu\text{g}/\text{ml}$ and 27 $\mu\text{g}/\text{ml}$, the monomers had approx. three- and sixfold higher specific activities than the aggregates respectively. The dependence of the initial rates of monomer self-phosphorylation on protein concentration was investigated in detail (Figure 5). Initial rates were measured over the first 10 s of the reaction, when the stoichiometry of self-phosphorylation was low (< 0.15 mol of phosphate/mol of protein; Figure 5b). The protein concentration dependence of the self-phosphorylation rate of the monomeric form was non-linear, increasing with enzyme concentration (Figure 5a). Within a 50-fold concentration range (0.1–5.0 μM), doubling the concentration of the monomeric cytoplasmic domain caused a modest 1.4-fold increase in catalytic rate. Logarithmic plots of the initial rate against protein concentration gave a slope of 1.5,

in agreement with a bimolecular cross-phosphorylation process (Figure 5c). The initial rate of self-phosphorylation of the aggregated form varied linearly with protein concentration (results not shown), as would be expected for a first-order process. Similar analysis showed that self-phosphorylation of the aggregated protein was a first-order reaction, which is consistent with intermolecular phosphorylation within aggregates (results not shown). Although these data are consistent with cross-phosphorylation, the rate increases resulting from raising protein concentration are lower than expected from an oligomerization-induced stimulation of catalytic activity.

The effects of Mn^{2+} , Mg^{2+} and $(\text{NH}_4)_2\text{SO}_4$ were investigated. Phosphorylation activity was stimulated strongly by Mn^{2+} and to a lesser extent by Mg^{2+} ; peptide phosphorylation activity was stimulated 2.4–3.5-fold at 10–20 mM Mn^{2+} relative to the activity at 1 mM Mn^{2+} , and 3.2–3.8-fold over the activity observed with Mg^{2+} at 10–20 mM; no attenuation of the activation was observed at high metal-ion concentrations.

Activation by $(\text{NH}_4)_2\text{SO}_4$

The cytoplasmic domain and truncated forms of it lacking the self-phosphorylation domain are activated by high concentrations of $(\text{NH}_4)_2\text{SO}_4$ towards phosphorylation of $[\text{Val}^5]$ angiotensin II (Wedegaertner and Gill, 1989, 1992) (K. Smith and M. Gregoriou, unpublished work). A 42 kDa calpain fragment of the cytoplasmic domain, lacking the self-phosphorylation domain and the basic juxtamembrane region sequence Arg⁶⁴⁵–Leu⁶⁵⁹, was purified and shown to be catalytically active in peptide- and self-phosphorylation assays (Gregoriou et al., 1994). $(\text{NH}_4)_2\text{SO}_4$ stimulated the activity of the 42 kDa fragment towards angiotensin II up to 15-fold (Figure 6a). Therefore the mechanism of $(\text{NH}_4)_2\text{SO}_4$ activation did not involve the N-terminal basic

Figure 7 Analysis of the $(\text{NH}_4)_2\text{SO}_4$ -stimulated self-phosphorylation by tryptic phosphopeptide mapping

(a) The effect of $(\text{NH}_4)_2\text{SO}_4$ and NaCl at high concentrations on the rate of self-phosphorylation of the cytoplasmic domain. Triplicate samples are shown in the autoradiographs under the corresponding salt concentrations used. (b) Cytoplasmic domain (5.2 mg of protein) was self-phosphorylated in the absence (i, ii) or presence (iii, iv) of 0.9 M $(\text{NH}_4)_2\text{SO}_4$ for 1 min (i, iii) or 18 min (ii, iv) on ice. (c) EGF receptor in A431 plasma membrane (200 mg of protein) was self-phosphorylated for 1 min on ice after preincubation with 0.8 M EGF for 10 min. Reactions were terminated by heating in Laemmli sample buffer, and proteins were purified, digested with trypsin and analysed as described in the Experimental section. ^{32}P -labeled phosphopeptide profiles were obtained by reverse-phase h.p.l.c. on a C_{18} column, and were developed with a linear gradient of acetonitrile in 0.1% (v/v) TFA. In the profiles shown, fractions 30 and 150 correspond to 3.84% and 38.4% (v/v) acetonitrile respectively.

sequences of the juxtamembrane domain. No activating effect was observed at low sulphate concentrations (50 mM), whereas both the intact and truncated domains required concentrations > 0.6 M for maximal activation. Activation of the 42 kDa fragment required higher concentrations of $(\text{NH}_4)_2\text{SO}_4$ (1.0–1.8 M) compared with the intact domain.

Activation of peptide phosphorylation by $(\text{NH}_4)_2\text{SO}_4$ has been shown to be substrate-specific (Hubler et al., 1992; McGlynn et al., 1992; Wedegaertner and Gill, 1992), while reports of the effect on self-phosphorylation appear to be conflicting (Wedegaertner and Gill, 1989; Hubler et al., 1992; Mohammadi et al., 1993). We investigated the substrate specificity of activation of the cytoplasmic domain by examining the effect of $(\text{NH}_4)_2\text{SO}_4$ on the phosphorylation of closely related peptide substrates and on self-phosphorylation. The rates of phosphorylation of both the src peptide RRLIEDAEYAARG and of band-3 peptide RREELQDDYEDD were decreased approx. twofold. To find out whether the nature of amino acid side chains at positions -1 or -2 relative to tyrosine in [Val⁵]angiotensin II (DRVYVHPF) might determine the specificity of activation, the effect of $(\text{NH}_4)_2\text{SO}_4$ on the phosphorylation of angiotensin-related peptides DREYVHPF and EGVYVHPV was tested (Figure 6b). The rates of phosphorylation of all three homologues of angiotensin II were essentially similarly enhanced by $(\text{NH}_4)_2\text{SO}_4$, indicating that the nature of amino acids at positions -1 and -2 by itself was not a crucial determinant of sulphate activation.

Rather, the strong hydrophobic character at positions $+1$ to $+4$ common to the angiotensin-related peptides may be involved in hydrophobic interactions with the active site, and would be promoted by high concentrations of $(\text{NH}_4)_2\text{SO}_4$.

The effect of $(\text{NH}_4)_2\text{SO}_4$ on self-phosphorylation was complex. At 0–0.25 M $(\text{NH}_4)_2\text{SO}_4$ self-phosphorylation was inhibited by $\sim 35\%$, whereas at 0.25–1.5 M the activity was increased up to threefold (Figure 6c). The initial inhibitory effect was mimicked by NaCl up to 1 M, and may therefore be attributed to ionic-strength effects. NaCl did not stimulate self-phosphorylation at ionic strengths corresponding to activating concentrations of $(\text{NH}_4)_2\text{SO}_4$. SDS/PAGE analysis showed that 1 M $(\text{NH}_4)_2\text{SO}_4$ led to formation of a highly phosphorylated kinase which appeared as a retarded major band of high radioactive content (Figure 7a). This contrasted with self-phosphorylation in the presence of 1 M NaCl, or in the absence of salt, and indicated that activating concentrations of $(\text{NH}_4)_2\text{SO}_4$ were associated with a highly phosphorylated state of the cytoplasmic domain.

We examined the possibility that sulphate might enhance catalytic activity by altering self-phosphorylation. Tryptic phosphopeptide maps of the cytoplasmic domain and of the EGF-stimulated whole receptor in A431 membranes were identical, the ^{32}P label being distributed similarly across the major peaks (Figure 7). Maps of the kinase domain after self-phosphorylation in the presence of $(\text{NH}_4)_2\text{SO}_4$ (0.9 M) and in its absence were compared [Figure 7b(i–iv)]. Although the

Figure 8 Far-u.v. c.d. of the monomeric cytoplasmic domain

The cytoplasmic domain was from Mono Q peak I, prepared by Procedure 1. (a) The dependence of the $[\theta]_{M,222}/[\theta]_{M,208}$ ratio on protein concentration, showing no significant conformational changes with increasing protein concentration. (b) Far-u.v. c.d. spectrum (●) of the protein (0.34 μM) in Tris/H₂SO₄ buffer, pH 7.5, at 10 °C, in the wavelength region 195–240 nm; the spectrum shows a good qualitative fit with a calculated curve (solid line) derived using the program CCA, as described in the Experimental section (Perczel et al., 1992). (c) The effect of (NH₄)₂SO₄ concentration on the far-u.v. c.d. spectrum of monomeric cytoplasmic domain under the same conditions; Concentrations of (NH₄)₂SO₄: ●, 0 M; ▲, 0.25 M; ■, 0.5 M; △, 0.75 M; and ○, 1 M. Similar spectra were obtained in three different experiments, using two different batches of protein and with (NH₄)₂SO₄ from two different suppliers.

phosphopeptide profiles obtained after self-phosphorylation for 1 min in the presence and in the absence of (NH₄)₂SO₄ were similar, two notable differences were observed. Extensive self-phosphorylation of virtually all phosphopeptides was achieved within 1 min of incubation with ATP in the presence of (NH₄)₂SO₄, peptides being phosphorylated to a larger extent than in its absence, even after prolonged incubation with ATP [Figure 7b(ii and iii)]. A minor peak (peak 7), was

significantly increased when (NH₄)₂SO₄ was present. Peak 7 was not observed after extensive (18 min) self-phosphorylation in the absence of salt. It therefore appears that (NH₄)₂SO₄ stimulates stable phosphorylation of additional sites which are not normally phosphorylated even after prolonged self-phosphorylation in its absence.

Secondary structure

The conformation of the monomeric cytoplasmic domain was probed using far-u.v. c.d., in order to define the major features of secondary structure, and to find out whether any protein–protein interactions could be detected by this technique. The far-u.v. c.d. spectra of the cytoplasmic domain were measured in the protein concentration range 0.17–1.7 μM in 50 mM Tris/H₂SO₄ (pH 7.5) at 10 °C. All the spectra had negative maxima at 208 and 222 nm and a positive maximum at 190 nm, similar to the standard helical spectra. With increasing protein concentration, the ratio of $[\theta]_M$ at 222 nm and 208 nm, $[\theta]_{M,222}/[\theta]_{M,208}$ did not appear to change appreciably (Figure 8a), indicating that no detectable changes in the conformation of the protein occurred with changing protein concentration over this range.

Secondary conformation was analysed by convex constraint analysis in terms of four pure conformational components, as described in the Experimental section (Perczel et al., 1992). These four components corresponded to α -helix (32%), β -sheet (6%), random coil (51%), and β -turn and other non-secondary structure contributions (10%). The observed c.d. spectra at 0.34 μM protein, and the calculated curve based on the reference protein set are displayed in Figure 8(b), which shows a good qualitative fit of the experimental data to the calculated curve. Therefore, in terms of secondary structure only, the cytoplasmic domain was estimated to contain approx. 35% α -helix, 58% random coil, and 7% β -sheet.

Effect of (NH₄)₂SO₄ on the c.d. spectra

The effect of (NH₄)₂SO₄ on the conformation of the cytoplasmic domain was also monitored by far-u.v. c.d. (Figure 8c). At 0.34 μM protein, at pH 7.5, little if any variation of the c.d. spectra occurred with increasing salt in the range 0–0.5 M. However, (NH₄)₂SO₄ at 0.75 M and 1 M caused a dramatic change of the shape of the c.d. spectra, with a decrease in ellipticity at 208 and 222 nm. It was estimated that at 1 M (NH₄)₂SO₄, both the random coil and α -helical content were reduced by ~50%. In contrast, there was an ~70% increase in the β -sheet content.

DISCUSSION

Kinase activation is important in EGF receptor signalling, hence many studies aim to characterize mechanisms responsible for increased kinase activity. Our initial preparations of soluble cytoplasmic domain had at least tenfold higher specific activity than those reported earlier (Wedegaertner and Gill, 1989). We therefore investigated preparation procedures, protein heterogeneity and oligomeric status as factors that might account for differences in the intrinsic activities of recombinant cytoplasmic domain preparations.

We overexpressed the cytoplasmic domain using the baculovirus transfer vector pAcYM1, and developed alternative procedures for isolation of milligram quantities of unphosphorylated, highly purified protein. The baculovirus expression system has been used to produce recombinant EGF

receptor and various constructs of the cytoplasmic domain (Greenfield et al., 1988; Wedegaertner and Gill, 1989; Hsu et al., 1990; McGlynn et al., 1992; Wedegaertner and Gill, 1992; Cadena et al., 1994). The preparations described here seemed to contain only trace amounts of phosphotyrosine, and were similar to native A431 EGF receptor with respect to the extent of self-phosphorylation *in vitro*. Our early preparations involving precipitation by $(\text{NH}_4)_2\text{SO}_4$ at 0.8 M, and those reported by others also using precipitation at 2 M and 1.5 M $(\text{NH}_4)_2\text{SO}_4$ (Hsu et al., 1990; McGlynn et al., 1992), had a tenfold higher turnover number than the immunopurified preparation, which required 1M $(\text{NH}_4)_2\text{SO}_4$ for full activity (Wedegaertner and Gill, 1989). In view of the activating effect of $(\text{NH}_4)_2\text{SO}_4$, we addressed the possibility that this purification step might cause activation of the kinase. We developed two similar but alternative purification protocols yielding protein of 95% purity, one of which included precipitation from crude extract at 0.8 M $(\text{NH}_4)_2\text{SO}_4$. Comparison of protein preparations performed simultaneously from the same cell batch showed that $(\text{NH}_4)_2\text{SO}_4$ did not increase irreversibly the intrinsic activity of the protein.

We investigated the oligomeric state of the cytoplasmic domain as a possible basis for its high specific activity. The protein was isolated as a monomer as shown by gel filtration and sucrose-density-gradient sedimentation. However, storage in pure form for 2–3 days resulted in the formation of soluble aggregates of lower specific activity, the proportion of which increased with storage time. The tendency of the protein to aggregate was spontaneous, irreversible, was dependent on time, temperature and protein concentration, and could be prevented by storage below 0 °C, where the protein remained monomeric for 1–2 years. This type of aggregation can therefore lead to variability in the activity of the cytoplasmic-domain preparations. Interestingly, aggregates of the native EGF receptor with either no activity or reduced activity have been reported (Biswas et al., 1985; Northwood and Davis, 1988). Purification of the cytoplasmic domain, by single-step anti-phosphotyrosine chromatography, was recently reported to yield a mixture of monomers and dimers of similar peptide phosphorylation activity, and aggregates of tenfold lower activity, which do not readily interconvert (Cadena et al., 1994), thus accounting for differences in specific activities of different preparations. Although our preparation of the cytoplasmic domain was homogeneously monomeric, heterogeneity was evident both in mass and charge, and did not appear to affect kinase activity.

In order to find out whether the cytoplasmic domain plays a role in receptor activation through oligomerization, we investigated the order of reaction of self-phosphorylation, by examining the dependence of initial rate on protein concentration. The results suggested that self-phosphorylation of monomers proceeded by cross-phosphorylation, as reported for detergent-solubilized monomeric receptor (Canals, 1992). The reaction order for the aggregated cytoplasmic domain was consistent with either interdomain phosphorylation within a domain aggregate, or with an intradomain mechanism. Reaction orders of 1 have been reported by Weber et al. (1984) for EGF-stimulated native receptor and by Yarden and Schlessinger (1987) for the unstimulated receptor. We found that the self-phosphorylation turnover number increased only 1.5-fold upon a twofold increase in the monomeric cytoplasmic-domain concentration, indicating that, under the assay conditions used, the monomeric cytoplasmic domain did not form activated oligomers. It is also noteworthy that in a similar protein concentration range, we found no evidence from c.d. for monomeric protein–protein interactions as protein concentration was increased tenfold. It has been suggested that an intermolecular mechanism applies for the

EGF-liganded receptor, and an intramolecular mechanism for the unliganded receptor (Yarden and Schlessinger, 1987). Our findings are thus consistent with an oligomerization/activation mechanism where the extracellular domain is required for receptor activation through stabilization of activated oligomers (Hurwitz et al., 1991).

High concentrations of $(\text{NH}_4)_2\text{SO}_4$ (1 M) and Mn^{2+} (10 mM) are known to stimulate the kinase activity of the EGF receptor and its cytoplasmic domain to levels comparable with those achieved by EGF (Pike et al., 1984; Wedegaertner and Gill, 1989; Koland and Cerione, 1990). Evidence has suggested that small ions might mediate their activating effects on the EGF receptor kinase in the monomeric state (Koland and Cerione, 1988), in contrast to the oligomerization-induced activation by EGF. However, recent data support aggregation as a mechanism of activation by $(\text{NH}_4)_2\text{SO}_4$ (Mohammadi et al., 1993). We found that $(\text{NH}_4)_2\text{SO}_4$ caused up to 13-fold stimulation of the $[\text{Val}^5]\text{angiotensin II}$ phosphorylating activity of both intact and truncated cytoplasmic domain. The effect of $(\text{NH}_4)_2\text{SO}_4$ on peptide phosphorylation was substrate-specific, as reported previously. Our results suggested that activation of peptide phosphorylation by $(\text{NH}_4)_2\text{SO}_4$ might be correlated with the presence of strong hydrophobic residues on the C-terminal side of the peptide phosphoacceptor tyrosine.

The effect on self-phosphorylation was more complex, being inhibitory in the range 0–0.25 M $(\text{NH}_4)_2\text{SO}_4$ and activating up to threefold at higher concentrations. Similar inhibition effects were also reported for the EGF receptor (Hubler et al., 1992), and our results suggest that it is likely to be mediated by ionic interactions. Further characterization of the activating effect of $(\text{NH}_4)_2\text{SO}_4$ on self-phosphorylation showed that the salt increased the rate and extent of phosphorylation, beyond that observed with the native EGF-stimulated receptor. An additional phosphopeptide was resolved in maps of $(\text{NH}_4)_2\text{SO}_4$ -stimulated cytoplasmic domain. It is not known whether this peptide corresponds to a novel site, and whether its phosphorylation causes activation of the kinase, as $(\text{NH}_4)_2\text{SO}_4$ causes a higher state of activation than the known activation level effected by EGF. Alternatively, the increase in the extent of self-phosphorylation is more likely to arise from increased exposure of tyrosine side chains caused by $(\text{NH}_4)_2\text{SO}_4$ -induced conformational changes in the cytoplasmic domain.

Far-u.v. c.d. analysis showed that the cytoplasmic domain had significant proportions of α -helical and random-coil structures. Using this technique, we found no evidence of conformational change that might occur through protein–protein interactions as protein concentration was increased; it is possible that such interactions might not be associated with conformational changes, but no evidence for rapid and spontaneous protein–protein interactions was found using alternative techniques, such as sedimentation analysis and gel filtration, both of which showed that the unphosphorylated protein was isolated in monomeric form. However, significant differences were detected between far-u.v. c.d. spectra obtained at 0 and 1 M $(\text{NH}_4)_2\text{SO}_4$, these differences were not seen between the spectra at 0 and 0.25 M salt, where self-phosphorylation activity is inhibited. Additional positive ellipticity at 208 nm was seen, which may arise from an increase in β -structure of some type (β -sheet or β -turn). Adoption of β -structure has often been observed to cause increased hydrophobicity of peptides, causing aggregation and precipitation (Hartman et al., 1974; Hilbich et al., 1991; Dyson et al., 1992; Yang et al., 1994). Maximum activation of the cytoplasmic domain, the calpain-generated 42 kDa catalytic fragment (Gregoriou et al., 1994), and a recombinant truncated protein lacking the autophosphorylation domain (K. Smith and M. Gregoriou, unpublished work) required levels of $(\text{NH}_4)_2\text{SO}_4$

which increased the hydrophobic character of the respective protein. It is possible that the observed increase in β -structure elicited by $(\text{NH}_4)_2\text{SO}_4$ (which may contribute significantly to aggregation and precipitation of the protein) is the basis for activation of its catalytic activity. Similar conformational changes may be involved in the activation of the cytoplasmic domain by other agents, the precipitating action of which has been recently correlated with their activating effects (Mohammadi et al., 1993). Kinetic evidence for a conformational basis for the potentiating effect of $(\text{NH}_4)_2\text{SO}_4$ on the EGF receptor has been reported (Hubler et al., 1992). Our observations suggest that the basis for kinase activation by $(\text{NH}_4)_2\text{SO}_4$ is a conformational change in the catalytic domain.

In summary, we report detailed alternative purification procedures for the isolation of active, monomeric, unphosphorylated EGF receptor cytoplasmic domain, and its physicochemical properties, including conformational studies by sucrose-density-gradient sedimentation and far-u.v. c.d. analyses. These preparations are suitable for structural and functional studies, in particular studies of structural changes induced by self-phosphorylation and oligomerization. Irrespective of treatment with activating concentrations of $(\text{NH}_4)_2\text{SO}_4$, the cytoplasmic domain was isolated as a uniformly monomeric protein that slowly and spontaneously aggregated into less-active soluble aggregates. Our results suggest that, whereas EGF-induced receptor oligomerization leads to kinase-domain activation, spontaneous association of the cytoplasmic domain, which lacks transmembrane and extracellular sequences, does not. These sequences may therefore play an important role in conferring specificity of activation through dimerization. Our investigation of the mechanism of activation by $(\text{NH}_4)_2\text{SO}_4$ suggests that conformational changes in the cytoplasmic domain of the EGF receptor may cause stimulation of its kinase activity.

We thank Glaxo Group Research, the Nuffield Foundation, the Edward Penley Abraham Research Fund, the MRC and the Oxford Centre for Molecular Sciences (OCMS) for funding the work of M.G., P.F.J., J.F.T. and J.J.Y. S.E.R. is a Royal Society 1983 University Research Fellow. We also thank Professor L. N. Johnson for useful discussions and support, Drs H. Overton, I. Jones and R. Possee (IVEM, Oxford) for vectors and guidance in baculovirus expression, Drs N. Hastie and R. Hill (MRC CAPCU, Edinburgh) for help with cloning, Dr T. A. Libermann for providing the v-erbB cDNA probe, Mr A. C. Willis (MRC Immunochimistry Unit, Oxford) and Dr G. Turcatti (GIMB, Geneva) for protein sequencing, Dr. Robin Aplin (OCMS, Oxford) for m.s., Dr K. Mofford (OCMS) for i.e.f., and Mr. Stephen Lee for photography.

REFERENCES

- Ackers, G. K. (1964) *Biochemistry* **3**, 723–730
- Beemon, K. and Hunter, T. (1978) *J. Virol.* **28**, 551–566
- Biswas, R., Basu, M., Sen, M. A. and Das, M. (1985) *Biochemistry* **24**, 3795–802
- Bohmer, F. D., Bohmer, S. A. and Heldin, C. H. (1993) *FEBS Lett.* **331**, 276–80
- Boni, S. M. and Pilch, P. F. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7832–7836
- Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254
- Braun, S., Ghany, M. A. and Racker, E. (1983) *Anal. Biochem.* **135**, 369–378
- Cadena, D. L., Chan, C. L. and Gill, G. N. (1994) *J. Biol. Chem.* **269**, 260–265
- Canals, F. (1992) *Biochemistry* **31**, 4493–4501
- Carter, P., Bedouelle, H. and Winter, G. (1985) *Nucleic Acids Res.* **13**, 4431–4443
- Chen, W. S., Lazar, C. S., Poenie, M., Tsien, R. Y., Gill, G. N. and Rosenfeld, M. G. (1987) *Nature (London)* **328**, 820–823
- Cooper, J. A., Sefton, B. M. and Hunter, A. T. (1983) *Methods Enzymol.* **99**, 387–403
- Downward, J., Parker, P. and Waterfield, M. D. (1984a) *Nature (London)* **311**, 483–485
- Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. and Waterfield, M. D. (1984b) *Nature (London)* **307**, 521–527
- Downward, J., Waterfield, M. D. and Parker, P. J. (1985) *J. Biol. Chem.* **260**, 14538–14546
- Dyson, H. J., Sayre, J., Merutka, G., Shin, H.-C., Lerner, R. A. and Wright, P. E. (1992) *J. Mol. Biol.* **226**, 819–835
- Erneux, C., Cohen, S. and Garbers, D. L. (1983) *J. Biol. Chem.* **258**, 4137–4142
- Fry, M., Panayotou, G., Booker, G. and Waterfield, M. (1993) *Protein Sci.* **2**, 1785–1797
- Fung, M.-C., Chiu, K. Y. M., Weber, T., Chang, T.-W. and Chang, N. T. (1988) *J. Virol. Methods* **19**, 33–42
- Greenfield, C., Patel, G., Clark, S., Jones, N. and Waterfield, M. D. (1988) *EMBO J.* **7**, 139–46
- Gregoriou, M. and Rees, A. R. (1984) *EMBO J.* **3**, 929–937
- Gregoriou, M., Willis, A. C., Pearson, M. A. and Crawford, C. (1994) *Eur. J. Biochem.* **223**, 455–464
- Hartman, R., Schwaner, R. C. and Hermans, J. (1974) *J. Mol. Biol.* **90**, 415–429
- Hilbich, C., Kisters-Woike, B., Reed, J., Masters, C. L. and Beyreuther, K. (1991) *J. Mol. Biol.* **218**, 149–163
- Honegger, A. M., Dull, T. J., Felder, S., Van, O. E., Bellot, F., Szapary, D., Schmidt, A., Ullrich, A. and Schlessinger, J. (1987) *Cell* **51**, 199–209
- Hsu, C. Y., Mohammadi, M., Nathan, M., Honegger, A., Ullrich, A., Schlessinger, J. and Hurwitz, D. R. (1990) *Cell Growth Differ.* **1**, 191–200
- Hsuan, J. J., Totty, N. and Waterfield, M. D. (1989) *Biochem. J.* **262**, 659–663
- Hubler, L., Kher, U. and Bertics, P. J. (1992) *Biochim. Biophys. Acta* **1133**, 307–315
- Hurwitz, D. R., Emanuel, S. L., Nathan, M. H., Sarver, N., Ullrich, A., Felder, S., Lax, I. and Schlessinger, J. (1991) *J. Biol. Chem.* **266**, 22035–22043
- Jones, P. F. (1987) D.Phil. Thesis, University of Oxford, Oxford
- Kemp, B. E. and Pearson, R. B. (1991) *Methods Enzymol.* **200**, 121–134
- Koland, J. G. and Cerione, R. A. (1988) *J. Biol. Chem.* **263**, 2230–2237
- Koland, J. G. and Cerione, R. A. (1990) *Biochim. Biophys. Acta* **1052**, 489–498
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Margolis, B. L., Lax, I., Kris, R., Dombalagian, M., Honegger, A. M., Howk, R., Givol, D., Ullrich, A. and Schlessinger, J. (1989) *J. Biol. Chem.* **264**, 10667–10671
- Martin, R. G. and Ames, B. N. (1961) *J. Biol. Chem.* **236**, 1372–1379
- Matsuura, Y., Possee, R. D., Overton, H. O. and Bishop, D. H. L. (1987) *J. Gen. Virol.* **68**, 1233–1250
- McGlynn, E., Becker, M., Mett, H., Reutener, S., Cozens, R. and Lydon, N. B. (1992) *Eur. J. Biochem.* **207**, 265–275
- Mohammadi, M., Honegger, A., Sorokin, A., Ullrich, A., Schlessinger, J. and Hurwitz, D. R. (1993) *Biochemistry* **32**, 8742–8748
- Northwood, I. C. and Davis, R. J. (1988) *J. Biol. Chem.* **263**, 7450–7453
- Overton, H. O., Ihara, T. and Bishop, D. H. L. (1987) *Virology* **157**, 338–350
- Perczel, A., Hollosi, M., Tusnadi, G. and Fasman, G. D. (1992) *Protein Eng.* **4**, 669–679
- Pike, L. J. (1987) *Methods Enzymol.* **146**, 353–362
- Pike, L. J., Kuenzel, E. A., Kasnellie, J. E. and Krebs, E. G. (1984) *J. Biol. Chem.* **259**, 9913–9921
- Siegel, L. M. and Monty, K. J. (1966) *Biochim. Biophys. Acta* **112**, 346–362
- Smith, M. H. (1970) in *Handbook of Biochemistry and Selected Data for Molecular Biology*, 2nd edn. (Sober, H. A., ed.), pp. C3–C281, CRC Press, Cleveland, OH
- Summers, M. D. and Smith, G. E. (1987) *Tex. Agric. Exp. Stn. Bull.* **1555**
- Thom, D., Powell, A., Lloyd, C. and Rees, D. (1977) *Biochem. J.* **168**, 187–194
- Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Downward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D. and Seeburg, P. H. (1984) *Nature (London)* **309**, 418–425
- Weber, W., Bertics, P. J. and Gill, G. N. (1984) *J. Biol. Chem.* **259**, 14631–14636
- Wedegaertner, P. B. and Gill, G. N. (1989) *J. Biol. Chem.* **264**, 11346–11353
- Wedegaertner, P. B. and Gill, G. N. (1992) *Arch. Biochem. Biophys.* **292**, 273–80
- Weyer, U., Knight, S. and Possee, R. D. (1990) *J. Gen. Virol.* **71**, 1525–1534
- Yang, J. J., Pitkeathy, M. and Radford, S. E. (1994) *Biochemistry*, in the press
- Yarden, Y. and Schlessinger, J. (1987) *Biochemistry* **26**, 1434–1442
- Yarden, Y. and Ullrich, A. (1988) *Ann. Rev. Biochem.* **57**, 443–478