Plant-based production of biopharmaceuticals
Rainer Fischer1,2, Eva Stoger1, Stefan Schillberg2, Paul Christou2 and Richard M Twyman3,4

Plants are now gaining widespread acceptance as a general platform for the large-scale production of recombinant proteins. The first plant-derived recombinant pharmaceutical proteins are reaching the final stages of clinical evaluation, and many more are in the development pipeline. Over the past two years, there have been some notable technological advances in this flourishing area of applied biotechnology, as shown by the continuing commercial development of novel plant-based expression platforms. There has also been significant success in tackling some of the limitations of plant bioreactors, such as low yields and inconsistent product quality, that have limited the approval of plant-derived pharmaceuticals.

Addresses
1Institute for Molecular Biotechnology, Biology VII, RWTH Aachen, Worringerweg 1, 52074 Aachen, Germany
2Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Grafschaft, Auf dem Aberg 1, 57392 Schmallenberg, Germany
3Department of Biology, University of York, Heslington, York YO10 5DD, UK
4e-mail: richard@writescience.com

Introduction
Proteins can be used as diagnostic reagents, vaccines and drugs, and this creates a strong demand for the production of recombinant proteins on an industrial scale. Commercial protein production has traditionally relied on microbial fermentation and mammalian cell lines, but these systems have disadvantages in terms of cost, scalability and safety that have prompted research into alternatives. Despite industry inertia and conservatism, plants have emerged as one of the most promising general production platforms for tomorrow’s biologics. Plants allow the cost-effective production of recombinant proteins on an agricultural scale, while eliminating risks of product contamination with endotoxins or human patho-

gens [1,2,3*,4*]. Another advantage of the use of plants in recombinant protein production is that vaccine candidates can be expressed in edible plant organs, allowing them to be administered as unprocessed or partially processed material [5].

Current limitations of plant bioreactor technology include the low yields that are achieved for many proteins (which are often caused by poor protein stability), difficulties with downstream processing (leading to inconsistent product quality), and the presence of non-authentic glycan structures on recombinant human proteins. These problems raise regulatory issues and have prevented the routine approval of plant-derived biopharmaceuticals for use in clinical trials [6,7]. In this review, we discuss technical advances achieved over the past two years that have helped to address these limitations, thus bringing the prospect of affordable, plant-derived biologics another step closer.

Emerging production platforms for biopharmaceuticals
Choosing a host species
Many of the early, plant-derived recombinant proteins were produced in transgenic tobacco plants and were extracted directly from harvested leaves. The continuing popularity of tobacco reflects its status as a well-established expression host for which robust transformation procedures and well-characterized regulatory elements for the control of transgene expression are available [1]. Furthermore, its high biomass yields and rapid scalability make tobacco very suitable for commercial molecular farming. It is also a non-food, non-feed crop, and so carries a reduced risk of transgenic material or recombinant proteins contaminating feed and human food chains [8]. Tobacco has been adopted as a platform system by several biotech companies, including Planet Biotechnology Inc. (http://www.planetbiotechnology.com/) and Meristem Therapeutics (http://www.meristem-therapeutics.com/), the only two companies to have plant-derived pharmaceuticals undergoing phase-II clinical trials.

One disadvantage of tobacco is its high content of nicotine and other toxic alkaloids, which must be removed completely during downstream processing steps. Although low-alkaloid tobacco cultivars are available, attention has turned to other leafy crops for pharmaceutical production. These crops include lettuce, which has been used for clinical trials with a hepatitis B virus subunit vaccine [9], and alfalfa, which is being promoted as a
platform system by Medicago Inc. (http://www.medicago.com). This Canadian biotech company has isolated novel promoters that allow high-level protein expression in alfalfa leaves, and has focussed on the early part of the production pipeline by developing alfalfa cell-culture and transient-expression technology [10]. Advantages of alfalfa include its high biomass yield and the fact that it is a perennial plant that fixes its own nitrogen. A strong advantage of alfalfa for pharmaceutical production is that fact that glycoproteins synthesised in alfalfa leaves tend to have homogeneous glycan structures, which is important for batch-to-batch consistency (see review by Gomord in this issue). However, alfalfa is a feed crop and its leaves contain large amounts of oxalic acid, which might interfere with processing.

Although leafy crops are advantageous in terms of biomass yield, proteins that are expressed in leaves tend to be unstable, which means the harvested material has a limited shelf life and must be processed immediately after harvest. By contrast, proteins that are expressed in cereal seeds are protected from proteolytic degradation; they can remain stable for up to three years at room temperature (E Stoger, unpublished data) and for at least three years at refrigerator temperature without significant loss of activity [11]. Several different cereals, including rice, wheat, barley and maize, have been investigated as expression hosts [8,12]. Maize has been chosen by Prodigene Inc. (http://www.prodigene.com), an industry leader in cereal-based commercial protein production, because it has a high biomass yield, because it is easily transformed and manipulated in vitro, and because the production of transgenic maize can be scaled up conveniently. Maize has been used for the commercial production of the technical proteins avidin and β-glucuronidase (GUS) [13,14]. In addition, Prodigene is exploring its use for the production of subunit vaccines [15], recombinant antibodies [16] and further technical enzymes, such as aprotinin and laccase [17].

Although it is beneficial to focus on a small number of platform technologies for the bulk production of biopharmaceuticals, the delivery of recombinant vaccines in edible plant organs is exceptional because it would be advantageous to use locally grown plants for vaccination campaigns. Therefore, a variety of different expression hosts have been evaluated. Potato was the first major system to be used for vaccine production, and transgenic potato tubers have been administered to humans in at least three clinical trials to date [18]. Over the past year, potatoes have been evaluated for the production of human serum albumin [19], novel vaccine candidates [20,21], tumour necrosis factor α (TNF-α) [22] and antibodies [23,24]. Other production hosts that have been used to express vaccines include tomatoes, bananas, carrots, lettuce, maize, alfalfa, white clover and Arabidopsis.

Oilcrops are useful hosts for protein production because the oil bodies can be exploited to simplify protein isolation. An example is the oleosin-fusion platform developed by SemBioSys Genetics Inc. (http://www.sembiosys.com/), in which the target recombinant protein is expressed in oilseed rape or safflower as a fusion with oleosin (see below; [P1]). The Finnish biotech company UniCrop (http://www.unicrop.fi/) is also developing an oilseed technology platform, although in this case, the idea is to isolate recombinant proteins from the rapidly developing sprouts cultivated in bioreactors.

Finally, there have been significant recent developments in the use of more diverse plant species, which can easily be contained, propagated and transformed, to produce recombinant proteins. Mayfield et al. [25**] have described a protein expression system that is based on the unicellular green alga Chlamydomonas reinhardtii. In this system, chloroplast-targeted transgenes were used to express an antibody that recognised herpes simplex virus glycoprotein D (see article by Franklin and Mayfield in this issue). Other simple plants that have been adopted as bioreactors include Lemna (duckweed), which is being developed as a platform technology by Biolex Inc. (http://www.biolex.com; [26]), and the moss Physcomitrella patens, which is being developed by Greenovation Inc., Freiburg, Germany (http://www.greenovation.com; see article by Decker and Reski in this issue). The advantages and disadvantages of different expression hosts are summarised in Table 1.

Alternative plant-based expression systems

The majority of plant-derived recombinant pharmaceutical proteins have been produced by nuclear transformation and the regeneration of transgenic plant lines, followed by the extraction and purification of proteins from the transgenic tissues. Although nuclear gene transfer is now routine in many species, it has disadvantages in terms of production time-scales, which are being addressed or circumvented by the development of alternative plant-based production technologies (Table 2).

Transient expression is generally used to verify transformation construct activity and to validate small amounts of recombinant protein. It can be achieved by the vacuum infiltration of leaves with recombinant Agrobacterium tumefaciens, resulting in the transient transformation of many cells [27]. High levels of protein expression are achieved for a short time, but generally the technique is insufficient for commercial-scale production [28,29]. Recently, however, several reports have described how this agroinfiltration process could be scaled-up more efficiently. Baulcombe and colleagues [30**] have shown that the loss of protein expression seen after a few days is predominantly caused by gene silencing. They managed to increase the expression levels of several proteins at least 50-fold by co-expressing the p19 protein from tomato
bushy stunt virus, a known inhibitor of gene silencing. Furthermore, researchers at Medicago Inc. have described how the agroinfiltration of alfalfa leaves can be scaled up to 7500 leaves per week, producing micrograms of recombinant protein each week [10]. Similarly, we have shown that up to 100 kg of wildtype tobacco leaves could be processed by agroinfiltration, resulting in the production of several hundred milligrams of protein (R Fischer, S Schillberg, unpublished).

Another emerging tobacco transient-expression technology is based on the use of plant viruses as expression vectors. Virus-infected plants have been used to produce several pharmaceutical proteins, including vaccine candidates and antibodies, one of which is now undergoing phase-I clinical trials [31]. The advantages of virus-based production include the rapid onset of expression, the systemic spread of the virus so that recombinant protein is produced in every cell, and the fact that more than one vector can be used in the same plant, allowing multimeric proteins to be assembled [32]. Plant virus expression systems are discussed by Gleba et al. in this issue.

The tobacco chloroplast transgenic system is another promising variant, which was boosted this year by the

### Table 1

<table>
<thead>
<tr>
<th>Species</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model plants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>Range of available mutants, accessible genetics, ease of transformation</td>
<td>Not useful for commercial production (low biomass)</td>
</tr>
<tr>
<td>Simple plants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physcomitrella patens,</td>
<td>Containment, clonal propagation, secretion into medium, regulatory compliance, homologous recombination in Physcomitrella</td>
<td>Scability</td>
</tr>
<tr>
<td>Chlamydomonas reinhardtii, Lemna</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leafy crops</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco</td>
<td>High yield, established transformation and expression technology, rapid scale-up, non-food/feed</td>
<td>Low protein stability in harvested material, presence of alkaloids</td>
</tr>
<tr>
<td>Alfalfa, clover</td>
<td>High yield, useful for animal vaccines, clonal propagation, homogenous N-glycans (alfalfa)</td>
<td>Low protein stability in harvested material, presence of oxalic acid</td>
</tr>
<tr>
<td>Lettuce</td>
<td>Edible, useful for human vaccines</td>
<td>Low protein stability in harvested material</td>
</tr>
<tr>
<td>Cereals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maize, rice</td>
<td>Protein stability during storage, high yield, easy to transform and manipulate</td>
<td>Low yields, difficult to transform and manipulate</td>
</tr>
<tr>
<td>Wheat, barley</td>
<td>Protein stability during storage</td>
<td></td>
</tr>
<tr>
<td>Legumes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soybean</td>
<td>Economical, high biomass, expression in seed coat</td>
<td>Low expression levels, difficult to transform and manipulate</td>
</tr>
<tr>
<td>Pea, pigeon pea</td>
<td>High protein content</td>
<td>Low expression levels</td>
</tr>
<tr>
<td>Fruits and vegetables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potato, carrot</td>
<td>Edible, proteins stable in storage tissues</td>
<td>Potato needs to be cooked</td>
</tr>
<tr>
<td>Tomato</td>
<td>Edible, containment in greenhouses</td>
<td>More expensive to grow, must be chilled after harvest</td>
</tr>
<tr>
<td>Oilsraps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oilseseed rape, Camelina sativa</td>
<td>Oleosin-fusion platform, sprouting system</td>
<td>Lower yields?</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>System</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transgenic plants, accumulation within plant</td>
<td>Yield, economy scalability, establishment of permanent lines</td>
<td>Production timescale, regulatory compliance</td>
</tr>
<tr>
<td>Transgenic plants, secretion from roots or leaves</td>
<td>Containment, purification</td>
<td>Scale, yield, cost of production facilities</td>
</tr>
<tr>
<td>Transplastomic plants</td>
<td>Yield, multiple gene expression, low toxicity, containment</td>
<td>Absence of glycosylation, some evidence of horizontal gene transfer</td>
</tr>
<tr>
<td>Virus-infected plants</td>
<td>Yield, timescale, mixed infections</td>
<td>Biosafety, construct-size limitations</td>
</tr>
<tr>
<td>Agroinfiltration</td>
<td>Timescale</td>
<td>Cost</td>
</tr>
<tr>
<td>Cell or tissue culture</td>
<td>Timescale, containment, secretion into medium (purification), regulatory compliance</td>
<td>Cost</td>
</tr>
</tbody>
</table>
Transplastomic plants are generated by introducing DNA into the chloroplast genome, usually by particle bombardment [33,34]. The advantages of chloroplast transformation are many: the transgene copy number is high because of the many chloroplasts in a typical photosynthetic cell, there is no gene silencing, multiple genes can be expressed in operons, the recombinant proteins accumulate within the chloroplast thus limiting toxicity to the host plant, and the absence of functional chloroplast DNA in the pollen of most crops provides natural transgene containment.

The chloroplast transgenic system has achieved remarkably high expression levels, recently exceeding 25% total soluble protein (TSP) for a tetanus toxin fragment [35*], 11% TSP for human serum albumen [36*] and 6% TSP for a thermostable xylanase [37*]. At present, chloroplast transformation is routine only in tobacco and *C. reinhardtii* (see above and the article by Franklin and Mayfield, this issue). However, plastid transformation has been achieved in a growing number of plant species, including carrot and tomato [33,34]. The ability to transform the chromoplasts of fruit and vegetable crops has obvious advantages for the expression of subunit vaccines [38].

Plant cell cultures can be used for the production of small-molecule drugs, but they are also advantageous for molecular farming because of the high level of containment that they offer and the possibility of producing proteins under current good manufacturing practice (cGMP) conditions [39]. Tobacco suspension cells are the most popular system at present, although pharmaceutical proteins have also been produced in soybean, tomato and rice cells, and in tobacco hairy roots [40–43]. More than 20 pharmaceutical proteins have been produced in plant cell-suspension cultures, including antibodies, interleukins, erythropoietin, human granulocyte-macrophage colony stimulating factor (hGM-CSF) and hepatitis B antigen [39]. Unfortunately, few of these proteins have been expressed at yields sufficient for commercial production. As discussed below, the problem of poor yields could be addressed in part by the use of optimised regulatory elements. For example, the expression of hGM-CSF in rice suspensions using an inducible promoter produced far greater yields than was possible using tobacco cells and a constitutive promoter [41].

**Strategies to improve protein yields**

The factors that affect recombinant protein yields in transgenic plants and other plant systems have recently been reviewed in detail [44]. The general approach is to maximise both the efficiency of all stages of gene expression and protein stability by appropriate subcellular targeting.

Transgene expression in plants used for molecular farming is often driven by the strongest available constitutive promoters. However, regulated promoters are increasingly used, particularly those that allow external regulation by physical or chemical stimuli [45]. Several novel inducible promoters that may be useful in molecular farming applications have been described recently. For example, a peroxidase gene promoter isolated from sweet potato (*Ipomoea batatas*) was used to drive the *gusA* reporter gene in transgenic tobacco. This promoter produced 30 times more GUS activity than did the cauliflower mosaic virus (CaMV) 35S promoter following exposure to hydrogen peroxide, wounding or ultraviolet light [46]. The wounding response is interesting as it would allow post-harvest induction of gene expression in the same manner as the CropTech mechanical gene activation (MeGA) system, which is based on a tomato hydroxy-3-methylglutaryl CoA reductase2 (HMGR2) promoter.

A novel seed-specific promoter from the common bean (*Phaseolus vulgaris*) has been used to express a single-chain antibody in *Arabidopsis thaliana*. In contrast to the CaMV 35S promoter, which resulted in antibody accumulation to 1% TSP, the bean *arc5-I* promoter resulted in antibody levels in excess of 36% TSP in homozygous seeds, and the antibody retained its antigen binding activity and affinity [47**]. A trichome-specific promoter that might be useful for the secretion of recombinant proteins into the leaf guttation fluid has also been described in tobacco [48]. Another secretion system, which is being commercialised by Phytomedics Inc. (http://www.phytomedics.com), involves the secretion of recombinant proteins into tobacco root exudates and the leaf guttation fluid. This was developed for the production of human secreted alkaline phosphatase and has recently been used for the secretion of recombinant antibodies [49*].

Subcellular targeting plays an important role in determining the yield of recombinant proteins because the compartment in which a recombinant protein accumulates strongly influences the interrelated processes of folding, assembly and post-translational modification. Comparative targeting experiments with full-size immunoglobulins and single-chain fragment variable (scFv) fragments have shown that the secretory pathway is more suitable for folding and assembly than the cytosol, and is therefore an advantageous site for high-level protein accumulation [50]. Antibodies that are targeted to the secretory pathway using either plant or animal amino-terminal signal peptides usually accumulate to levels that are several orders of magnitude greater than those of antibodies expressed in the cytosol. Occasional exceptions to this general observation suggest that the intrinsic features of each antibody might also contribute to overall stability [51]. The endoplasmic reticulum (ER) provides an oxidising environment and an abundance of molecular chaperones...
but few proteases. These features are likely to be the most important factors affecting protein folding and assembly. It has been shown recently that antibodies that are targeted to the secretory pathway in transgenic plants interact specifically with the molecular chaperone BiP [52].

In the absence of further targeting information, proteins that accumulate in the secretory system are secreted to the apoplast. Depending on its size, the protein can be retained in the apoplast or might leak from the cell, with important implications for production systems that are based on cell-suspension cultures. The stability of antibodies in the apoplast is lower than that in the lumen of the ER. Therefore, antibody expression levels can be increased even further if the protein is retrieved to the ER lumen using an H/KDEL carboxy-terminal tetrapeptide tag [53]. Accumulation levels of proteins tagged in this way are generally 2–10-fold greater than those of identical proteins that lack the KDEL signal [44]. ER retention can also influence the structure of glycan chains on plant-derived proteins ([54,55] and our unpublished observations), but we do not discuss the post-translational modification of plant-derived proteins any further because this topic is treated in detail by Gomord and Faye in this issue. Targeting is especially important if the recombinant protein is toxic to the production host. For example, the accumulation of avidin in the cytosol of transgenic tobacco plants is toxic, but plants can be regenerated successfully when this molecule is targeted to the vacuole [56].

**Downstream processing**

Although high-level expression is necessary to provide good yields in plant-based production systems, the efficient recovery of recombinant proteins must also be optimised. Secretion systems are advantageous because no disruption of plant cells is necessary during protein recovery; hence, the release of phenolic compounds is avoided. Nevertheless, the recombinant proteins may be unstable in the culture medium. The use of affinity tags to facilitate the recovery of proteins is a useful strategy as long as the tag can be removed after purification to restore the native structure of the protein. In the olesin fusion system mentioned earlier, the fusion protein can be recovered from oil bodies using a simple extraction procedure and the recombinant protein separated from its fusion partner by endoprotease digestion [P1]. Similarly, we have devised a strategy in which recombinant proteins are expressed as fusion constructs that contain an integral membrane-spanning domain derived from the human T-cell receptor, and are then purified from membrane fractions [57]. Recent strategies that have been described include the expression of His-tagged GUS-fusion proteins in tobacco chloroplasts [58], the extraction of His-tagged proteins by foam fractionation [59], and the release of recombinant proteins using a modified intein expression system [60•].

**Conclusions**

Plants have many advantages over established production technologies for the large-scale expression of recombinant proteins, but several challenges remain to be addressed in terms of improving yields and product quality. A small number of plant-derived biologics are approaching commercialisation, but these are the minority that have met the technological challenges, cleared the regulatory hurdles and overcome inertia in the biotechnology industry. We are facing a growing demand for protein therapeutics and diagnostics, but the capacity to meet those demands using established facilities is lacking. A shift to plant bioreactors might therefore become necessary within the next few years, making it more imperative that these issues are addressed and solved.

**Acknowledgements**

The authors would like to thank their colleagues at the Fraunhofer IME, RWTH Aachen and other institutes for their contribution to the field of molecular farming. Funding from the European Community project PharmaPlant is gratefully acknowledged.

**References and recommended reading**

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest


The authors describe the production of a large single-chain antibody in the chloroplasts of Chlamydomonas reinhardtii, and demonstrate that the antibody forms dimers, accumulates in a soluble form, and binds to its target antigen (herpes simplex virus glycoprotein D) in enzyme-linked immunosorbent assays (ELISAs).


The authors show that co-expression of a target protein and the p19 protein from tomato bushy stunt virus, a known suppressor of post-transcriptional gene silencing, results in a greater than 50-fold increase in the level of target proteins produced in agroinfiltrated leaves. The target proteins could be purified from as little as 100 mg of leaf tissue.


A novel seed-specific antigen-promoter, and the antibody remained functional and able to bind its target to levels nearly 40-fold higher than those achieved with the CaMV 35S promoter, and the antibody remained functional and able to bind its target antigen.

Yields of up to 100,000 molecules, were successfully secreted from the roots of transgenic tobacco plants growing in hydroponic culture medium. Yields of up to 11.7 μg antibody per gram of dry root mass per day were achieved.

This report is notable because monoclonal antibodies, which are large molecules, were successfully secreted from the roots of transgenic tobacco plants growing in hydroponic culture medium. Yields of up to 11.7 μg antibody per gram of dry root mass per day were achieved.

In this work, transgenic tobacco plants were used to produce a mammalian antimicrobial peptide called SMAP-29. The peptide was expressed as a fusion with a modified VMA intein expression system. In this work, transgenic tobacco plants were used to produce a mammalian antimicrobial peptide called SMAP-29. The peptide was expressed as a fusion with a modified VMA intein expression system. Purification of the peptide was achieved by exploiting the self-cleaving activity of the intein. Correct processing was demonstrated and the purified peptide retained its antimicrobial activity.

Function and glycosylation of plant-derived antiviral monoclonal antibody. Proc Natl Acad Sci USA 2003, 100:8013-8018.

Patent