High level expression of carotenogenic genes for enzyme purification and biochemical characterization

Gerhard Sandmann

Biosynthesis Group, Botanical Institute, J.W. Goethe Universität, P.O. Box 111932, D-60054 Frankfurt, Germany

Abstract: Heterologous expression in Escherichia coli of several carotenogenic enzymes is described as a successful strategy for biochemical characterization and elucidation of enzyme mechanisms. This approach is applicable for many different carotenogenic enzymes from bacteria, fungi and plants. Information is presented on the types of expression constructs used and the strategies employed for expression. Furthermore, details on purification of expressed proteins and on their enzymatic assays are given together with a survey on the biochemical properties of all carotenogenic enzymes which were obtained by expression in E. coli.

INTRODUCTION

For use as food colorants, animal feed and for other applications, an increasing number of carotenoids is isolated from natural sources or produced biotechnologically by super-producing or transgenic organisms (ref. 1). A broad knowledge about the function of the enzymes involved is important for a better understanding of the biosynthesis steps and the regulation of the carotenogenic pathway. Little information was available on the mechanism of carotenogenic enzymes until recently. Problems hampering isolation, purification and subsequent biochemical characterization of carotenogenic enzymes were, besides their mostly membrane-bound nature, their low abundance in cells and tissue and their sensitivity towards detergents when they are solubilized (ref. 2). Nevertheless, it was possible to purify some carotenogenic enzymes as active preparations, namely phytoene synthase (ref. 3), phytoene desaturase (ref. 4) and capsanthin synthase (ref. 5) from chromoplasts of Capsicum fruit. possible due to the very high biosynthetic capacity of this organelle and the fact that it can be easily isolated. Attempts to purify carotenogenic enzymes from other material were less successful. Therefore, we present our experience with a new strategy for the biochemical characterization of homogenous carotenogenic enzymes by use of heterologous gene expression (refs. 6,7) to overcome the quantitative limitations and to facilitate the purification of the corresponding enzyme. The recent progress in cloning of different genes of the carotenogenic pathway from bacteria, fungi and higher plants provides the tools for this approach. Furthermore, a survey on the biochemical information about carotenogenic enzymes obtained by this technique to date will be presented.

PROTEIN EXPRESSION

A number of expression systems is available including the bacteria *Escherichia coli* or *Bacillus subtilis* and also eukaryotic ones like insect cells and the yeast *Pichia pastoris* (ref. 8). We started our expression work with the best-established *E. coli* system for which a wide range of different expression vectors is available. Until now we were able to overexpress, to an amount of 1 to 10% of total *E. coli* cell protein, any carotenogenic gene which we tried, regardless whether this was of prokaryotic or eukaryotic origin. All the purified and characterized enzymes dealt with in this report were expressed in *E. coli*.

Expression constructs

The expression vectors used in general were the multicopy pUC plasmids which are available with extensive multiple cloning sites inserted into the 5' region of the lacZ gene and in different orientations (ref. 9). In general, protein synthesis is under control of the lacZ promoter and can be stimulated by IPTG, but because of the high copy number of the plasmid not enough synthesized repressor protein is available

2164 G. SANDMANN

for complete suppression. For example, we needed a tight control of protein expression in the case of crtB from Erwinia. Otherwise the expressed phytoene synthase was lethal to E. coli because it drained off all the geranylgeranyl diphosphate (GGDP) necessary for other metabolic reactions. This happened when we used our pUC-derived expression vector alone. To avoid this, we co-transformed with the compatible plasmid pREP4 (from Quiagen Inc.), a multicopy repressor plasmid which ensures high levels of lac repressor, resulting in tight regulation of protein expression. Then protein expression was started at the end of the growth phase by addition of IPTG. Many of our expressed enzymes had N-terminal extension of the polypeptide chain of up to 16 amino acids due to the cloning procedures. We never found any indication that this affected enzymatic activity. However, it should be pointed out that, for maximum activity of the enzymes, the transit sequence present in eukaryotic cDNAs must be deleted when making the expression construct (ref. 10). The E. coli strain used routinely by us for expression with pUC vectors is JM101. Recently, we changed to pQE vectors (Quiagen Inc.) which behave similarly to the pUC family but have the advantage that they additionally encode for a 6 histidine tag, in our case at the N-terminal end of the protein. This is a very useful affinity group to aid purification of the enzymes by chromatography on resins with Ni-nitrilotriacetic acid ligands. This type of purification of native enzymes worked quite well in our hands with phytoene desaturase from Neurospora crassa and ζ-carotene desaturase from Capsicum but failed with the Erwinia uredovora phytoene synthase. Obviously, the histidine tag is not accessible and is hidden due to the folding of the protein in this case.

Another very useful cloning vector is pT7-7 (ref. 11). It contains the T7 RNA polymerase promoter which gives extremely strong expression. We used it in a case in which expression of an enzyme was comparatively low with pUC. Another advantage of pT7-7 is its NdeI restriction site. Cloning into this site results in an expression construct which starts with its own ATG avoiding any N-terminal extention of the expressed protein. For expression with this vector, *E. coli* strain BL21(DE3) should be the host. It contains the T7 RNA polymerase gene integrated into its genome behind the lacZ promoter which makes the system IPTG inducible.

A problem may occur with *E. coli* as expression host in the presence of unfavorable codons for this bacterium. This is often the case for genes from GC rich organisms. But we also had a similar problem with a carotenogenic gene from *E. uredovora*. The first N-terminal amino acid leucine was encoded by TTG which is quite rare for *E. coli*. This was the reason why expression was extremely low. After replacing in the expression plasmid this codon for leucine by CTG which is a more favoured one, protein expression was as high as expected.

Expression strategies

Most of the carotenogenic enzymes are membrane-bound. Upon expression, the major part of a synthesized enzyme is sequestered in inclusion bodies and only a minor portion is integrated in *E. coli* membranes where it is active. Starting from these inclusion bodies facilitates the purification of the enzyme because all other proteins can easily be removed by detergent-washing steps. For subsequent solubilization of the inclusion body protein we used 8 M urea. At the end of the purification process, an inactive enzyme is obtained. The major effort is then the renaturation by refolding of the protein. We could successfully reactivate two different phytoene desaturases from *E. uredovora* (ref. 12) and *Synechococcus* (ref. 13). Reactivation was achieved in the first case by removal of urea and incubation for 30 min. with 10 mM dithiothreitol. The same basic procedure was applied to reactivate the phytoene desaturase from *Synechococcus*. However, reactivation of the enzyme was improved more than two-fold upon addition of lipids. The most favorable were monogalactosyl diacylglycerol and phospholipid.

Alternatively to inclusion bodies, the portion of the expressed enzyme which is integrated into the membranes can be isolated. In one case we were able to solubilize a carotenogenic enzyme from the membrane with detergents without loss of activity. This was the ζ -carotene desaturase from *Anabaena* (ref. 14). The detergents of choice were Chaps and Nonidet. With many other detergents which we tried, activity was not or was only poorly retained.

As some carotenogenic enzymes like the *E. uredovora* lycopene cyclase (ref. 15) are very unstable in the presence of all available detergents, we developed a procedure to express membrane-bound carotenogenic enzymes in a way that French press homogenization of *E. coli* cells gives some of the enzyme in a soluble but nevertheless active form. The conditions used involve growth of the transgenic *E. coli* culture at moderate temperatures at about 28°C and cell breaking with a pressure of up to 90 MPa. This was used for the expression of phytoene desaturase from *Rhodobacter capsulatus* (ref. 16) and the lycopene cyclase from *E. uredovora* (ref. 15). However, it should be mentioned that with other enzymes this very high pressure destroyed enzyme activity.

PROTEIN PURIFICATION

With highly expressed carotenogenic enzymes purification can easily be monitored from the beginning by SDS polyacrylamide gel electrophoresis. The enzyme can be recognized by the prominent band which is absent in a non-induced sample or in a control transformant carrying the cloning vector without an inserted reading frame. Due to the high concentration of expressed enzyme compared to the other accompanying proteins, only one or two purification steps are sufficient to obtain a homogeneous protein. Proteins expressed with a 6-His tag can be purified by affinity chromatography in one step after optimization of the binding to Ni-NTA and removal of unspecifically-binding proteins by different washing conditions. For proteins without an affinity group, chromatography on DEAE cellulose was the most efficient chromatographic step. In the case of GGDP synthase from E uredovora (ref. 17) and phytoene desaturase from E capsulatus (ref. 16) this step was sufficient to obtain homogenous enzymes. For purification of the E-carotene desaturase from Anabaena a second step involving hydrophobic interaction chromatography on pentyl agarose was indispensable (ref. 14).

An unconventional way of purification had to be developed for lycopene cyclase from *E. uredovora* (ref. 15). This enzyme is highly lipophilic. After French press treatment of *E. coli* cells and centrifugation it is found in the supernatant in an active form from which it can be precipitated by 40% ammonium sulphate. However after this step, it could not be redissolved in buffer again. In addition, solubilization by use of detergents was not feasible because, with any available detergent, enzyme activity was destroyed. Thus, we washed away contaminating proteins from the ammonium sulphate precipitate with buffer and were then able to extract lycopene cyclase with a lipid-containing buffer. Minor contaminating proteins were subsequently removed by immuno absorption. For this purpose the lipid-extracted fraction was passed through a column filled with CNBr-agarose to which an antiserum against total *E. coli* proteins was coupled.

ENZYME ASSAYS

Due to the low activities found in preparations from non-transformed organisms, *in vitro* assays of carotenogenic enzymes were, until recently, only possible with radioactive-labelled substrates which had to be synthesized enzymatically from ^{14}C -mevalonic acid or isopentenyl diphosphate. This limited enzymatic studies because only a few ^{14}C -carotenes could be obtained e.g. by using fungal mutants (ref. 18). With all our overexpressed enzymes activity was high enough to avoid the use of radioactive substrates. This allows the application of many different carotenoids when substrate specificity of an enzyme is investigated. Usually, an amount of about 1 μ g of carotene is sufficient as substrate and gives enough reaction product for the determinaton by HPLC. But there is still the problem of how to make a lipophilic hydrocarbon accessible to a mainly membrane-bound enzyme in an aqueous system. Two ways are possible. If a carotenogenic enzyme is stable and active in the presence of a detergent, the substrate carotenes can be added to the reaction mixture containing 0.1% of this detergent as an acetone solution. When detergents must be avoided, we routinely make suspensions from soybean lipids to which we apply the carotenoids. In the reaction mixture the lipids mediate the transfer of their carotenoid load to the enzyme. We used this way of substrate application for a phytoene desaturase (ref. 16), a ζ -carotene desaturase (ref. 14) and a lycopene cyclase (ref. 15).

PROPERTIES OF ENZYMES

There are several indications that carotenoid biosynthesis relies on an independent synthesis and supply of geranylgeranyl diphosphate (GGDP). For example, the *crtE* gene necessary for GGDP synthesis is part of the carotenogenic gene clusters in all bacteria investigated to date (ref. 2). Furthermore, during fruit ripening which is accompanied by massive carotenoid formation in *Capsicum*, expression of the GGDP synthase gene is strongly enhanced (ref. 19). The same is observed for the GGDP synthase gene in *Neurospora crassa* during light regulation of carotenoid biosynthesis (ref. 20). Consequently, properties of GGDP synthase are of considerable interest in studies of the carotenoid biosynthetic pathway.

To date only enzymes involved in the reaction sequence from synthesis of GGDP via synthesis of phytoene, desaturations and cyclization to β-carotene have been expressed, purified and biochemically characterized. Their properties, which are summarized in Table 1, will be described in the following paragraphs.

2166 G. SANDMANN

TABLE 1. Biochemical characteristics of expressed and purified enzymes of the carotenoid biosynthetic pathway

Enzyme	Source	Substrate*	Product	Cofactors	Inhibitors	Ref.
GGDP-Syn.	Erwinia	GDP (9 μM) FDP (11 μM)	GGDP GGDP			(17)
Phyt-Syn.	Erwinia	GGDP (14 μM)	Phytoene	ATP	Phosphate, Squalestatin	(23)
Phyt-Desat.						
Δ2	Synechococcus	Phytoene (3.5 μM) 1,2-Epoxyphytoene	ζ-Carotene Epoxy-ζ-carotene	NAD(P) (14 mM)	Norflurazon etc.	(13)
Δ3	Rhodobacter	Phytoene (33.3 μM) Phytofluene ζ-Carotene (16.6 μM)	Neurosporene Neurosporene Neurosporene	FAD (7 μM)	Diphenylamine	(16)
$\Delta 4$	Erwinia	Phytoene	Lycopene	FAD	Diphenylamine	(12)
Δ5	Neurospora	Phytoene	Lycopene 3,4-Dehydro- lycopene			
ζ-Car-Desat.						
crtQ	Anabaena	ζ-Carotene (9.7 μM) Neurosporene (10.3 μM) β-Zeacarotene Pro-ζ-carotene	Lycopene Lycopene γ-Carotene Prolycopene	Oxygen		(14)
zds	Capsicum	ζ-Carotene	Neurosporene Lycopene			
Lycop-Cycl.	Erwinia	Lycopene (1.8 μM) γ-Carotene Neurosporene (6.3 μM)	B-Carotene B-Carotene B-Zeacarotene	NAD(P)H (2.5 mM)	СРТА	(15)

^{*}Km values are given in parenthesis.

GGDP synthase

Prenyl transferases add one or more isopentenyl diphosphate (IDP) molecules to an allylic acceptor. Among them are the GGDP synthases which are defined by their reaction product. Depending on the organisms, the allylic substrate of the reaction can either be the C5 dimethylallyl diphosphate (DMADP) which is condensed with three molecules of IDP, the C10 geranyl diphosphate reacting with two molecules of IDP, or the C15 farnesyl diphosphate with one IDP. In all reactions the C20 GGDP is formed (ref. 2). Our major interest in the purified GGDP synthase from E. uredovora was on its chain length specificity. Km values were determined for IDP, DMADP, GDP and FDPP (Table 1). These Km values together with the reaction rate indicated that GDP as well as FDP are the genuine allylic substrates for the GGDP synthase from E. uredovora. Product analysis demonstrated that GGDP was the only reaction product. In comparison other GGDP synthases from carotenogenic organisms, like the GGDP synthase from Phycomyces, are monospecific for FDP (ref. 21) whereas Capsicum GGDP synthase converts all allylic diphosphates from C5 to C15 equally well (ref. 22).

 $[\]Delta$ indicates the number of desaturation steps carried out by the different phytoene desaturases.

Phytoene synthase,

Phytoene is formed by the condensation of two molecules of GGDP with prephytoene diphosphate as an intermediate. This reaction, leading to the synthesis of phytoene, is carried out by a single enzyme. We expressed the phytoene synthase from E uredovora (ref. 23). Its properties, which were obtained with a partially purified preparation, were compared to those of the plant enzyme from Capsicum (ref. 3). With a K_m value of 41 μ M for GGDP affinity for this substrate is about 10 times lower for the bacterial than for the Capsicum enzyme. The E uredovora phytoene synthase was stimulated by ATP as co-substrate together with a divalent metal ion like Mg^{2+} or Mn^{2+} . Both enzymes were inhibited by inorganic phosphate. Furthermore, with the expressed enzyme from E uredovora we could demonstrate inhibition by squalestatin, an inhibitor of squalene synthase (ref. 24). The type of inhibition was non-competitive with respect to GGDP.

Desaturases

β-Carotene and derivatives are synthesized in bacteria, fungi and plants (ref. 2). Their biosynthesis involves desaturation of phytoene to lycopene, which contains four additional double bonds. From the molecular genetic work and the biochemical studies in recent years it is evident that basically different enzymes are involved in this desaturation sequence. Based on sequence similarity, two groups of carotenoid desaturases can be distinguished (ref. 25). The <u>crtI type</u> includes the bacterial and fungal phytoene desaturases, hydroxyneurosporene desaturase and the ζ-carotene desaturase CRTQ (formerly called ZDS) from *Anabaena*. To the <u>pds type</u> belong the phytoene desaturases from cyanobacteria and plants as well as the ζ-carotene desaturase (ZDS) from *Capsicum* (ref. 26). Biochemically, the crtI type and pds type of desaturases differ by the cofactors used as hydrogen acceptor during the desaturation reaction and by their sensitivity to inhibitors. Studies on substrate utilization and product formation of several expressed and purified crtI type phytoene desaturases revealed that enzymes exist which can carry out maximally 3, 4 or 5 desaturation steps. In Table 1 they are indicated as Δ3, Δ4 and Δ5 desaturases.

crtI types

The phytoene desaturase from R capsulatus is the one among this group that carries out the least number of desaturation steps. The purified enzyme introduces a maximum of three double bonds ($\Delta 3$) and converts phytoene or the intermediates phytofluene and ζ -carotene into neurosporene (ref. 16). Another substrate which can be desaturated by this enzyme is 1,2-epoxyphytoene yielding 1,2-epoxyneurosporene as the product. However, the C_{30} diapophytoene was not converted. The K_m values for the substrates phytoene and ζ -carotene were 33.3 and 16.6 μ M, respectively. Desaturation by this enzyme is dependent on only FAD as cofactor with a K_m value of 4.9 μ M. Other compounds like oxidized nicotine nucleotides or ATP had no effect. The I_{50} value for inhibition by diphenylamine was determined as 7 μ M. The purified phytoene desaturase from E uredovora inserts 4 double bonds into phytoene (ref. 12). This enzyme also uses FAD as a cofactor and is inhibited by diphenylamine. Both enzymes mentioned are closely related to the phytoene desaturase from N crassa. The principal difference of the fungal enzyme is the maximal number of 5 double bonds introduced into phytoene yielding 3,4-dehydrolycopene as reaction product.

The amino acid sequence of ζ -carotene desaturase from *Anabaena* exhibits similarities to the crtI type phytoene desaturases described above. However, this enzyme from *Anabaena* is unable to convert either phytoene or phytofluene. The genuine substrate is ζ -carotene which is converted to lycopene with a K_m value of 9.7 μ M, but also the intermediate neurosporene (K_m value 10.3 μ M) is desaturated to lycopene (ref. 14). In addition, β -zeacarotene and the di-cis pro- ζ -carotene are used as substrates. No need for a coenzyme was found for this enzyme but it was observed that the reaction is oxygen-dependent.

pds types

All pds type desaturases which we expressed and purified carry out two desaturation steps. In case of the phytoene desaturase from *Synechococcus*, phytoene and also 1,2-epoxyphytoene are converted to ζ -carotene or its epoxide (ref. 13). The cofactors involved are NAD or NADP and the reaction is inhibited by typical bleaching herbicides like norflurazon. The expressed ζ -carotene desaturase from *Capsicum* converts ζ -carotene to lycopene via neurosporene.

Lycopene cyclase

Lycopene can be cyclized to form β or ϵ rings. The β -cyclase from E. uredovora is the only lycopene cyclase purified and characterized to date (ref. 15). Cyclization carried out by this enzyme is a two-step reaction with the monocyclic γ -carotene as an intermediate. Other substrates for this enzyme were neurosporene and 1-hydroxylycopene which yielded β -zeacarotene and 1'-hydroxyl γ -carotene, respectively. Substrates which were not converted include 1,1'-dihydroxylycopene and the tetra-cis prolycopene. The cyclization of all substrates is dependent on NADH or NADPH as cofactors. Very

2168 G. SANDMANN

similar to that of higher plants, this bacterial lycopene cyclase exhibits the same inhibition properties towards substituted alkyl amines like CPTA. A comparable study on the enzymological properties of a higher plant lycopene cyclase will show how closely it resembles the bacterial enzyme and may help to decide whether a different enzyme was aquired in cyanobacteria and plants.

OUTLOOK

Due to the enormous progress in cloning of carotenoid genes, many carotenogenic enzymes can now be purified and studied after heterologous expression. The technique of expression could in our hands be universally applied to about 10 different carotenogenic enzymes from bacteria, fungi, algae and plants. Yields of purified protein are in general high enough not only for biochemical investigations but also to attempt structure elucidation after crystallization and by NMR studies. Of course, the ability to express enzymes, e.g. from the plant xanthophyll pathway, is limited by the availability of carotenogenic genes. But this will be possible soon, especially as the gene for zeaxanthin epoxidase from *Nicotiana plumbaginifolia* has been cloned very recently (ref. 27). It should be pointed out that bacteria are more diverse in synthesizing carotenoids than any other group of organisms. From carotenogenic species with unique pathways novel genes can be cloned in the near future which will make it possible to express and study other enzymes involved in xanthophyll biosynthesis. This will increase our knowledge on various modification reactions of carotenoids and how different branches of the carotenoid biosynthetic pathway operate.

ACKNOWLEDGEMENTS

The author thanks all his collaborators who were involved in the development of the techniques described here and in characterization of carotenogenic enzymes. This work was supported by the Deutsche Forschungsgemeinschaft and by a grant from the European Communities Biotech Programme, as part of the Project of Technological Priority 1993-1996.

REFERENCES

- 1. A. M. Nonomura. In: Carotenoids: Chemistry and Biology (N. I. Krinsky, M. M. Mathews-Roth and R. F. Taylor, eds.) pp. 365-375, Plenum Press, New York (1989). G. Sandmann. Eur. J. Biochem. 223, 7-24 (1994)
- 3. O. Dogbo, A. Laferriere, A. d'Harlingue, and B. Camara. Proc. Natl. Acad. Sci. USA 85, 7054-7058
- 4. P. Hugueney, S. Römer, M. Kuntz and B. Camara. Eur. J. Biochem. 209, 399-407 (1992)
- F. Bouvier, P. Hugueney, A. d'Harlingue, M. Kuntz and B. Camara. *Plant J.* 6, 45-54 (1994).
 R. C. Hockney. *TIBTECH* 12, 456-463 (1994).
- 7. G. Gross and H. Hauser. J. Biotechnol. 41, 91-110 (1995).
- 8. W. B. Frommer and O. Ninnemann. Annu. Rev. Plant Physiol. Plant Mol. Biol. 46, 419-444 (1995).
- 9. Z. Hanna, C. Fregeau, G. Prefonteine and R. Brousseau. Gene 30, 247-250 (1984).
- 10. N. Misawa, M. R. Truesdale, G. Sandmann, P. D. Fraser, C. Bird, W. Schuch and P. M. Bramley. J. Biochem. 116, 980-985 (1994).
- 11. S. Tabor and C. C. Richardson. Proc. Natl. Acad. Sci. USA 262, 1074-1078 (1985).
- 12. P. D. Fraser, N. Misawa, H. Linden, S. Yamano, K. Kobayashi and G. Sandmann. J. Biol. Chem. 267, 19891-19895 (1992)
- 13. P. D. Fraser, H. Linden and G. Sandmann. Biochem. J. 291, 687-692 (1993)
- 14. M. Albrecht, H. Linden and G. Sandmann. Eur. J. Biochem. 236, 115-120 (1996).

- G. Schnurr, N. Misawa and G. Sandmann. Biochem. J. 315, 869-874 (1996).
 A. Raisig, G. Bartley, P. Scolnik and G. Sandmann. J. Biochem. 119, 559-564 (1996).
 M. Wiedemann, N. Misawa and G. Sandmann. Arch. Biochem. Biophys. 306, 152-157 (1993).
- P. D. Fraser, M. Albrecht and G. Sandmann. J. Chromatogr. 645, 265-272 (1993).
 M. Kuntz, S. Römer, C. Suire, P. Huegeney, J. H. Weil, R. Schantz and B. Camara. Plant J. 2, 25-34 (1992).
 S. Baima, G. Macino and G. Morelli. J. Photochem. Photobiol. B: Biol. 11, 107-115 (1991).

- F. Lütke-Brinkhaus and H. C. Rilling, Arch. Biochem. Biophys. 266, 607-612 (1988).
 B. Camara, J. Bousquet, C. Cheniclet, J. P. Cardet, M. Kuntz, J. L. Evrar and J. H. Weil. In: Physiology, Biochemistry and Genetics of Non-green Plastids (C. D. Boyer, J. C. Shanon and R. C. Hardison, eds.) pp. 141-156, American Society of Plant Physiology, Washington DC, (1989).
 23. U. Neudert. Thesis Universität Frankfurt (1995).
 24. B. Cox, J. L. Hutson, S. E. Keeling, B. E. Kirk, A. R. P. Srikantha and N. S. Watson. Bioorg. Medic. Chem. Lett. 4, 1024 (1994).
- 1931-1936 (1994).
- 25. G. Sandmann. J. Plant Physiol. 143, 444-447 (1994).
- 26. M. Albrecht, A. Klein, P. Hugueney, G. Sandmann and M. Kuntz. FEBS Lett. 372, 199-202 (1995). 27. E. Marin, L. Nussaume, A. Quesada, M. Gonneau, B. Sotta, P. Hugueney, A. Frey and A. Marion-Poll. EMBO J. 15, 2331-2342 (1996).