CMP-N-acetylneuraminic acid synthetase of *Escherichia coli*: high level expression, purification and use in the enzymatic synthesis of CMP-N-acetylneuraminic acid and CMP-neuraminic acid derivatives

Spencer L. Shames¹, Ethan S. Simon¹², C. William Christopher, Walther Schmid¹, George M. Whitesides¹ and Li-Li Yang

Genzyme Corporation, Cambridge, MA 02139 and ¹Department of Chemistry, Harvard University, Cambridge, MA 02138, USA
²Present address: Biotage Inc., Charlottesville, VA 22901, USA

To whom correspondence should be addressed

The gene encoding CMP-N-acetylneuraminic acid (CMP-NeuAc) synthetase (EC 2.7.7.43) in *Escherichia coli* serotype O7 K1 was isolated and overexpressed in *E. coli* W3110. Maximum expression of 8–10% of the soluble *E. coli* protein was achieved by placing the gene with an engineered 5'-terminus and Shine-Dalgarno sequence into a pKK223 vector derivative behind the tac promoter. The overexpressed synthetase was purified to >95% homogeneity in a single step by chromatography on high titre Orange A Matrex™ dye resin. Enzyme purified by this method was used directly for the synthesis of CMP-NeuAc and derivatives. The enzymatic synthesis of CMP-NeuAc was carried out on a multigram scale using equimolar CTP and N-acetylneuraminic acid as substrates. The resultant CMP-NeuAc, isolated as its disodium salt by ethanol precipitation, was prepared in an overall yield of 94% and was judged to be >95% pure by 1H NMR analysis. N-Carboxymethoxyneuraminic acid and N-carbobenzyloxyneuraminic acid were also found to be substrates of the enzyme; S-azidoneuraminic acid was not a substrate of the enzyme. N-Carboxymethoxyneuraminic acid was coupled to CMP at a rate similar to that observed with NeuAc, whereas N-carbobenzyloxyneuraminic acid was coupled >100-fold more slowly. The high level of expression achieved with the *E. coli* synthetase, together with the high degree of purity readily obtainable from crude cell extracts, make the recombinant bacterial enzyme the preferred catalyst for the enzymatic synthesis of CMP-N-acetylneuraminic acid.

Key words: CMP-N-acetylneuraminic acid/CMP-N-acetylneuraminic acid synthetase/enzymatic synthesis/sialylation

Introduction

Sialylated biomolecules are found widely in nature and have been implicated in a variety of biorecognition processes. The structures of naturally occurring sialylated species range from the polymeric forms of bacterial colominic acid to the various glycolipids and glycoproteins found in animals and man. Recent interest in sialylated biomolecules has been fuelled by studies on host cell recognition by a variety of mammalian pathogens including influenza A and B viruses (Suzuki et al., 1986, 1987), Sendai virus (Suzuki et al., 1983, 1984) and the bacterium *Mycoplasma pneumoniae* (Loomes et al., 1984). The preparation of suitable substrates for the study of these host–pathogen interactive processes requires the isolation of naturally occurring sialylated biomolecules or, alternatively, the synthetic sialylation of glycoproteins and glycolipids.

The biosynthesis of sialylated biomolecules is dependent on two enzymatic reactions:

i) Sialic acid + CTP → CMP-sialic acid + PP

ii) CMP-sialic acid + glycosyl-OH → glycosyl-sialic acid + CMP

The first reaction, catalysed by CMP-N-acetylneuraminic acid (CMP-NeuAc) synthetase, activates NeuAc at the expense of one CTP equivalent. The second reaction, catalysed by a family of CMP-sialic acid glycosyltransferases, transfers sialic acid from the activated species to targeted glycosyl moieties. In similar fashion to the biological process, these two enzymes have been used for the synthetic preparation of a variety of sialylated glycoproteins (Rearick et al., 1979; Higa and Paulson, 1985; Gross et al., 1987; Gross and Brossmer, 1988) and glycolipids (Rearick et al., 1979). The most common source of CMP-NeuAc synthetase used for enzymic CMP-NeuAc synthesis has been bovine brain. The enzyme prepared from this source, however, is relatively impure and contains a high level of phosphatase activity that significantly reduces the reaction yield (Higa and Paulson, 1985; Auge and Gautheron, 1988; Simon et al., 1988).

It was of interest to determine if an overexpressed CMP-NeuAc synthetase, isolated from *Escherichia coli*, would provide a superior enzyme preparation for the enzymic synthesis of activated sialic acid derivatives. The CMP-NeuAc synthetase encoding gene from *E. coli* has been isolated (Vann et al., 1987) and sequenced (Zapata et al., 1989) previously. Using an oligonucleotide probe based on the preliminary sequence data of Vann et al. (1987), we have also obtained the synthetase encoding gene. This report describes the high level expression of CMP-NeuAc synthetase in *E. coli*, a convenient assay to monitor synthetase activity, a simple single-step column purification of the overexpressed enzyme and a protocol for use of the enzyme in the multigram scale synthesis of CMP-NeuAc.

Results and discussion

Overexpression of CMP-NeuAc synthetase

The gene encoding CMP-NeuAc synthetase of *E. coli* was isolated as described under Materials and methods. Initial expression experiments with the cloned CMP-NeuAc synthetase encoding gene were carried out by isolating the gene on a 2.7 kb *EcoRI/HindIII* fragment from pCMP.3 and ligating the fragment into the complementary sites on vector pKK223-NaeI behind the tac promoter (Figure 1). This
construct, given the name pKKCMP.1, was used to transform *E. coli* W3110. Induction of the transformed cells with 2.0 mM isopropl-β-D-thiogalactoside (IPTG) yielded crude extracts with optimal CMP-NeuAc expression at ~1.8 U/mg of the soluble protein. To increase the yield of enzyme production, the synthetase encoding DNA insert was truncated at the 3'-terminus and both the 5'-terminus and Shine-Dalgarno sequences were engineered. Truncation was carried out by digesting the 2.7 kb *EcoRI*/*HindIII* fragment with *NsiI* that cleaved 112 bp 3'-proximal to the synthetase termination codon. The resulting *EcoRI*/*NsiI* synthetase encoding fragment, now ~1.4 kb in length, was ligated into the *EcoRI*/*PstI* sites of vector pKK223-NacΔ to yield vector pKKCMP.2. Induction of *E. coli* W3110 harbouring pKKCMP.2 showed an overall ~2-fold increase in synthetase expression level over that obtained with the pKKCMP.1 vector. In a final engineering experiment, the parent Shine-Dalgarno sequence was modified to yield a more extended ribosome binding site with a stretch of eight adenine residues 5'-proximal to the initiation codon. This extended ribosome binding site has been demonstrated to provide high efficiency in *E. coli* translational initiation (Sullivan et al., 1989; Gold and Strom, 1990). To carry out this manipulation, a single silent mutation was made in construct pKKCMP.2 to convert the unique *XhoII* site located 40 bp from the initiation codon to a unique *BamHI* site. The synthetase Shine-Dalgarno sequence and 40 bp of the 5'-terminus were subsequently removed by digestion with *EcoRI* and *BamHI*. The N-terminal encoding gene fragment with the desired Shine-Dalgarno sequence was then replaced by utilizing a 62 bp synthetic oligonucleotide cassette. In the course of this manipulation, several codons were modified to those expected in genes encoding highly expressed *E. coli* proteins (Gouy and Gautier, 1982). The sequence of the engineered ribosome binding site and 5'-terminus of the synthetase encoding gene is shown in Figure 2. This final construct, given the designation pKKCMP.SD, gave expression levels ~5-fold higher than that achieved with vector pKKCMP.1 and resulted in typical specific activities of 0.25 U/mg in crude cell extracts of induced cells. This value is ca. 833-fold higher than that reported previously for expression of the isolated CMP-NeuAc synthetase encoding gene (Zapata et al., 1989).

### Purification of CMP-NeuAc synthetase

Recombinant CMP-NeuAc synthetase was purified ~11.2-fold from crude cell extracts of *E. coli* in an overall yield of 50–55%. Purification to >95% homogeneity, as judged by Coomassie blue staining of a 9.5% SDS–polyacrylamide gel (Figure 3), was routinely carried out by single-step column chromatography of crude cell extracts on high titre Orange A Matrex™ dye resin followed by ammonium sulphate and precipitation. The majority of contaminants were removed in the chromatographic procedure by extensive washing of the Orange A Matrex™ dye resin. The CMP-NeuAc synthetase was purified further by elution from the dye resin with a linear KCl gradient. A typical protein elution profile from the Orange A Matrex™ dye column is shown in Figure 4. A slight increase in enzyme purity was subsequently achieved by fractionation of the pooled chromatographic fractions with 60% ammonium sulphate.

A specific activity of 2.5–2.8 U/mg was routinely obtained for the purified recombinant CMP-NeuAc synthetase. This value is comparable to the specific activity of 2.1 U/mg reported previously by Vann et al. (1987) for enzyme purified from a wild type *E. coli* strain. Minor contaminants remaining in the enzyme preparation may be removed by affinity chromatography on CDP-hexanolamine arrose (data not shown). However, for preparing enzyme to be used in the synthesis of CMP-NeuAc, this latter purification step was unnecessary.

The purification of overexpressed CMP-NeuAc synthetase from *E. coli* by chromatography on Orange A Matrex™ resin provides an efficacious method of enzyme recovery from crude cell extracts. Successful purifications were carried out when Orange A resin with a dye content of 3.77 and 4.22 mg/ml was employed. However, Orange A resin with 2.3 mg dye/ml only bound the synthetase weakly and was unsatisfactory for purification. The yield of purified synthetase of 10.7 U/50 g of *E. coli* W3110 transformed with vector pKKCMP.SD is >200-fold higher than that obtained from an equivalent quantity of bovine brain tissue.

---

### Table 1. Construction of CMP-NeuAc synthetase expression vectors.

<table>
<thead>
<tr>
<th>Vector</th>
<th>Insert Restriction Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCMP 3</td>
<td>HindIII EcoRI XhoII NsiI NsiI HindIII</td>
</tr>
<tr>
<td>pKKCMP 1</td>
<td>EcoRI XhoII NsiI NsiI HindIII</td>
</tr>
<tr>
<td>pKKCMP 2</td>
<td>EcoRI XhoII NsiI NsiI HindIII</td>
</tr>
<tr>
<td>pKKCMP SD</td>
<td>EcoRI BamHI NsiI NsiI HindIII</td>
</tr>
</tbody>
</table>

**Fig. 1.** Construction of CMP-NeuAc synthetase expression vectors. A graphical representation of the steps carried out to enhance the expression of the CMP-NeuAc synthetase encoding gene is shown. The *E. coli* CMP-NeuAc synthetase structural gene is represented by the stippled box and non-coding DNA as solid lines. The tac promoter lies 5'-proximal to the *EcoRI* site in vector pKK223-NacΔ.

### Table 2. Synthetase sequence.

<table>
<thead>
<tr>
<th>Synthetic sequence:</th>
<th>Engineered sequence:</th>
</tr>
</thead>
<tbody>
<tr>
<td>AATTCAAAAAATCATCCGCGGAATA ATG AGA ACA AAA ATT ATG GCC GTA AT AAT CCA GCC CGT AGT G...</td>
<td>AATTCAATAGCGAATCCCGAAAA AAA ATG CGT ACT AAG ATC TTC GCC ATC CGC CGT AGT G...</td>
</tr>
</tbody>
</table>

**Fig. 2.** Sequence of the native and engineered ribosome binding sites and 5'-terminus of the CMP-NeuAc synthetase encoding gene. The Shine-Dalgarno sequences are underlined.
The multigram scale synthesis of CN4P-NeuAc was carried out successfully in ~94% yield with the purified recombinant CMP-NeuAc synthetase as catalyst. The enzyme, as purified from E.coli, showed no detectable phosphatase or nucleotidase activity as evidenced by the stoichiometric conversion of CTP to the corresponding CMP-NeuAc derivative. This can be compared to syntheses that required the addition of 3 mole equivalents of CTP to drive the reaction to completion when the partially purified calf brain enzyme was used (Simon et al., 1988). Furthermore, the high yield of purified synthetase readily obtained from 1 l of bacterial culture allows for the addition of more catalyst to the reaction mixture and, thus, decreases both the reaction time and the extent of spontaneous CTP hydrolysis.

The CMP-NeuAc obtained from the preparative synthetic reaction using the recombinant E.coli enzyme displayed 1H (500 MHz) and 31P (200 MHz) NMR spectral data identical to that obtained with authentic material from Sigma, as well as material prepared previously using the bovine brain enzyme (Simon et al., 1988). Purity was judged to be > 95%, by the lack of additional resonances above the baseline of the high field spectra.

**Alternative substrates of CMP-NeuAc synthetase**

To begin to examine the substrate specificity of E.coli CMP-NeuAc synthetase, three sialic acid derivatives, 5-N-carbo- benzoxyl-, 5-N-carbomethoxy- and 5-azidoneuraminic acid were evaluated as CMP acceptors. As shown in Table I, the carbomethoxy and carbobenzyloxy derivatives of neuraminic acid were alternative substrates of the E.coli synthetase, whereas no reaction was detected with the 5-azido derivative. At a concentration of 5.0 mM, 5-N-carbomethoxyneuraminic acid was turned over at a rate equal to that observed with NeuAc and >100-fold faster than the rate observed with 5-N-Cbz-neuraminic acid. Because of limiting quantities of the NeuAc analogues, no further kinetic characterization was carried out. To determine if the enzymic synthesis of the CMP-NeuAc analogues could be carried out on a larger scale and to verify that the observed rates of P, formation in the kinetic assays were indicative of CMP coupling, the synthetase reactions were carried out on a 0.1 mmol scale and monitored by 1H NMR. Whereas the N-carbomethoxy derivative showed >80%, product formation in 4 h and complete coupling in 14 h, the formation of CMP-Cbz-neuraminic acid only proceeded to ~10%, conversion after 22 h. Thus, the relative rates observed in the preparative scale reactions were consistent with those found with the kinetic assay.

**Table I. Substrate specificity of CMP-NeuAc synthetase of E.coli**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Relative rate</th>
<th>CMP adduct</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2CC(O)NH-</td>
<td>1</td>
<td>&gt; 90%, 4 h</td>
</tr>
<tr>
<td>H2CO(O)NH-</td>
<td>1</td>
<td>&gt; 90%, 4 h</td>
</tr>
<tr>
<td>Cbz-NH-</td>
<td>&lt; 0.01</td>
<td>&lt; 10%, 22 h</td>
</tr>
<tr>
<td>N5-</td>
<td>Not detectable</td>
<td>None detected</td>
</tr>
</tbody>
</table>

* Substitution at the C-5 position of sialic acid. Cbz = C6H5CH2CO(OH)-.
* Relative rate determined at 5.0 mM substrate under standard assay conditions.
* CMP adduct detected by 1H NMR spectroscopy as described in the text. Values entered are estimates of percent conversion at given times.
evolve by examining similar 9-substituted sialic acid derivatives as substrates of the E.coli enzyme.

Conclusion

The enzymic synthesis of CMP-NeuAc from NeuAc and CTP is operationally simpler using recombinant CMP-NeuAc synthetase isolated from E.coli in comparison to using the enzyme isolated from calf brain for three reasons. First, the single-step isolation of overexpressed CMP-NeuAc synthetase from bacterial culture requires less effort than the multi-step procedure for isolating the enzyme from calf brain. Secondly, the absence of contaminating phosphatase activity in the E.coli preparation eliminates the requirement for excess CTP in the enzymic reaction mixture. Lastly, the absence of side products caused by contaminating enzyme activities in the calf brain preparation simplifies the isolation of CMP-NeuAc: a single ethanol precipitation step produces CMP-NeuAc of sufficient purity for use in subsequent enzyme-catalyzed reactions.

Materials and methods

Inorganic pyrophosphatase and CTP were from Sigma. V-Acetylneuraminic acid was from either Sigma or Genzyme. Restriction enzymes and T4 DNA ligase were from Bethesda Research Laboratories. Calf intestine alkaline phosphatase was from Boehringer-Mannheim. Orange A Matrix* dye resins were from Amicon. SDS-polyacrylamide gel electrophoresis molecular weight standards were from BioRad. All other biochemical reagents were the highest quality available from Sigma and Aldrich, unless otherwise indicated.

Neuraminic acid derivatives

The synthetics of 5-N-carboxymethoxy-, 5-N-carboxbenzoxyl- and 5-acidone neuraminic acids were carried out from V.A growth medium (4,2 mg d/lum, was from Bethesda Research Laboratories. All strains were grown at 37°C on 2% TTY medium that contained 16 g Bactotryptone, 10 g Bacto yeast extract and 5 g NaCl/l of culture medium. Growth medium was supplemented with 75 µg ampicillin/ml when required. For enzyme preparations, bacterial cultures were grown to an OD660 of 1.2-1.4 and induced by adjusting the medium to 2.0 mM IPTG. Cells were harvested by centrifugation following a 5-h induction period. Cloning vector pUC19 was from Bethesda Research Laboratories. Vector pK223-Nae3 was prepared by complete digestion of vector pK223-3 (Pharmacia) with NaeI, followed by gel purification and ligation with T4 DNA ligase.

DNA manipulations and sequencing

All manipulations of DNA were carried out using standard techniques based on those described by Maniatis et al. (1982) or by manufacturers' instructions. Sequencing was performed by the dideoxy method of Sanger et al. (1977) according to the directions of the Sequenase kit. Oligonucleotides were prepared using an Applied Biosystems 380-B DNA synthesizer using standard phosphoramidite chemistry. The resultant oligonucleotides were purified by preparative polyacrylamide gel electrophoresis. When desired, oligonucleotides were phosphorylated on the 5' end with T4 polynucleotide kinase prior to annealing. Oligonucleotides were annealed in 77 mM Tris buffer (pH 7.5), containing 1 mM MgCl2 and 1 mM dithiothreitol (DTT) by heating to 85°C for 10 min, followed by slow cooling to room temperature.
The authors would like to thank Ms. Kathy Hahn and Mr. Gary White for the synthesis and purification of the oligosaccharides used in this work, and Dr. Andreas Schrell for the gift of the various C-5 substituted sialic acid derivatives.

Abbreviations
bp, base pair; DTT, dithiothreitol; IPTG, isopropyl β-D-thiogalactoside; kb, kilobase pair; NeuAc, N-acetyllactosamine; SDS, sodium dodecyl sulphate.

References

Received on November 1, 1990; accepted on November 29, 1990.