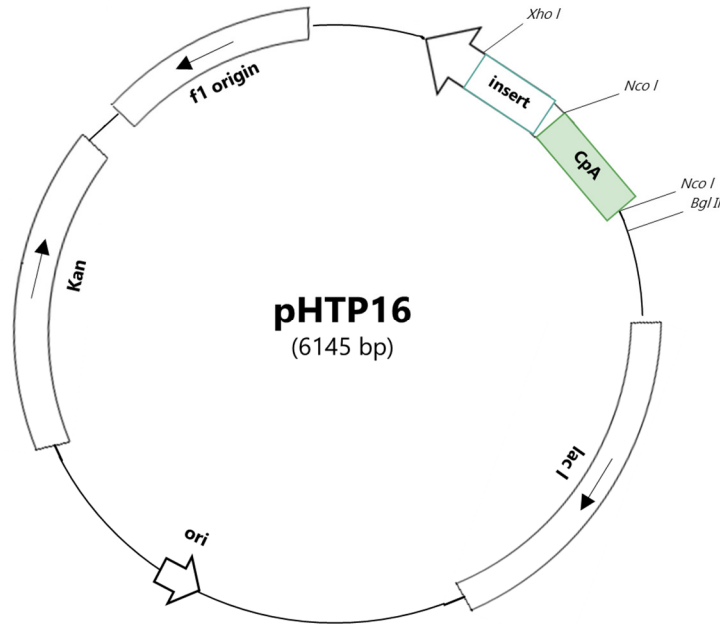


## pHTP16 Expression Vector

pHTP16 was designed for the cloning and expression of high-levels of recombinant proteins in *Escherichia coli*. Recombinant proteins are expressed in fusion with a *Ruminococcus flavefaciens* cellulosomal protein (CpA), which is able to promote solubility and folding of fusion partners. This vector, included in the portfolio of NZYTech pHTP expression vectors, is part of the NZYEasy Cloning & Expression System. pHTP16 contains two poly-histidine (6xHis) sequences (N- and C-terminal) which allow subsequent recombinant protein purification by immobilized metal ion affinity chromatography (IMAC).

### 1. Vector Map



#### pHTP16 Cloning/Expression Region

<i>Nco I</i>	CpA	<i>Nco I</i>	His-Tag
<u>CCATGGG</u> TGGCGAGTGTACGGCTATATCGAC.804bp..GATGTCAAGGATTACTCAATCGTTG		<u>CCATGGG</u> GCAGCAGCCATCATCATCATCACAGCAGCGGC	
MetGlyGlyGluCysHisGlyThrIleAsp.268aa..AspValLysAspTyrSerIleValAlaMetGlySerSer		HisHisHisHisHisHisSerSerGly	
CCTCAGCAAGGGCTGAGG />	CCTCAGCTCCGCTGAGGTCCGTCGACAAGCTTGC	<u>CTCGAGC</u> ACCACCACCACCACCAC	His-Tag
ProGlnGlnGlyLeuArg />	ProGlnLeuProLeuArgSerValAspLysLeuAlaAlaAlaLeuGlu	HisHisHisHisHisHis*	STOP

> Represents the site where the gene will be inserted.

**Note 1:** For correct expression, inserted gene needs to be in frame with pHTP16 5' gene sequence. Inserts correctly cloned into pHTP16 will maintain reading frames starting on the ATG codon.

**Note 2:** CpA is a recombinant cellulosomal protein (Cp) that is highly expressed in *E. coli*. This protein has carbohydrate-binding module activity, displaying affinity for  $\beta$ -glycans (xyloglucan, glucomannan, galactomannan and barley  $\beta$ -glucan).

## 2. Vector Sequence (6145 bp)

TGGCGAATGGGACGCGCCCTGTAGCGCGCATTAAAGCGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCCTACACTTGCAGCGCCCTAGCGCCCGCTCTTTTCGCTTTCTCCCT  
 TCCTTTCTCGCCACGTTTCGCGCGCTTTCCCGCTCAAGCTCTAAATCGGGGGCTCCCTTAGGGTTCGATTAGTGTCTTACGGCACCTCGACCCAAAAAATTGATTAGGGTGATG  
 GTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTGACGTTGGAGTCCACGTTCTTAAATAGTGGACTCTTGTCCAAACTGGAACACACTCAACCCATATCTCGGT  
 CTATTTCTTTGATTATAAGGGATTTTCCGGATTTCGGCTATTGGTTAAAAAATGAGTGTATTAAACAAAAATTAACCGGAATTTAAACAAAAATTAACGTTTACAATTTCAAGT  
 GGCACCTTTTCGGGAAATGTGCGCGGAACCCCTATTGGTTTATTTCCTAATACATTCAAATATGTATCCGCTCATGAATTAATCTTAGAAAAAATCATCGACATCAATGAAAC  
 TGCAATTTATTCATATCAGGATTATCAATACCATATTTTGA AAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCAGGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTCTG  
 CGATTCGGACTCGTCCAACTCAATACAACCTATTAATTTCCCTCGTCAAAAAAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAAATGGCAAAAGTTT  
 ATGCATTTCTTTCCAGACTTGTTCACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAACCCGTTATTCATTCTGTGATTGGCCTGAGCGAGACGAAATACCGCA  
 TCGCTGTAAAAAGGACAATTAACAACAGGAATCGAATGCAACCGCGCAGGAACTGACGAGCATCAACAATATTTTACCTGAATCAGGATATTTCTTAATACCTGGAATGCTG  
 TTTTCGGGGATCGCAGTGGTGAATTAATCGCGCCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTA  
 AACATATTGGCAACGCTACCTTTGGCATGTTTTCAGAAACACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTGCGACCTGATTGCCCGACATTATCGCGAGCCCATTTA  
 TACCCATATAAAATCAGACTCCATGTTGGAATTTAATCGCGCCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTA  
 TTGTTCAATACAAAAATCCCTTTTCGTTCCACTGAGCGTACAGCCCGTAGAAATGATCAAAAGGATCTTCTTGAGATCCCTTTTCTGCGGTAATCTGCTGCTTGC  
 AAACAAAAAACACCCGCTACCAGCGGTGTTGTTTTCGGGATCAAGAGCTACCAACTCTTTTCCGAAAGGTAACCTGGCTTACGACAGCGCAGATACCAAAATACCTGCTTCTAGT  
 GTAGCGGTAGTTAGGCCACCACTCAAGAACTCTGTAGCACCGCTACATACCTCGTCTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG  
 GACTCAAGAGGATAGTTACCGGATAAGCGCAGCGGTGCGGCTCAAGCGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGCACTGAGATACCTACAGCTGAGC  
 TAGGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGTATCCGGTAAAGCGGAGGTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTA  
 TAGTCTGTGCGGTTTCGCCACTCTGACTTTGAGCGTCAATTTTGTGATGCTCGTCAAGGGGGCGAGCCTATGGA AAAACCGCAGCAACCGCGCCCTTTTACGGTTCCCTGGCCTTT  
 TGCTGGCTTTTGCTCATGTTCTTCTCGGTTATCCCTGATTCTGTGGATAACCGTATTACCGCTTTGAGTGAGCTGATACCGCTCGCCGACGGCAACGACCGAGCGCAGCG  
 AGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGATTTTCTCCTTACGCATCTGTGCGTATTTCACACCGCATATATGGTGACTCTCAGTACAATCTGCTCTGATGCCCA  
 TAGTTAAGCCAGTATACACTCCGCTACGCTACGTTGGTCTGGTTCGCGCCCGACCCCGCAACACCCGCTGACGCGCCCTGACGGGCTGTCTGCTCCCGGCTCCCGCTA  
 CAGCAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAAGGTTTTACCGTCAATACCGAAACCGCGCAGGAGCAGTGGCGTAAGGCTCATCAGCCTGTGTGAAGCGATTCCACAG  
 ATGTCTGCCTGTTCATCCGCTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTCTGTTTGGTCACTGATGCT  
 CCGTGTAAAGGGGATTTCTGTTTCATGGGGTAAATGATACCGATGAAACGAGAGAGGATGCTCAGGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACTGTGAGGGTAA  
 ACAACTGGCGGTATGGATCGCGGGGACAGAAAAATCACTCAGGTCATGCCAGCGCTCTGTTAATACAGATGTAGGTGTTCCACAGGTTAGCCAGCAGCATCTGCGATGCG  
 ATCCGGCAATAATGGTGCAGGCTGACTTCCGCTTTCCAGACTTTACGAAACACGGAACCGAAGCAATTCATGTTGTTGCTCAGGTCGACAGACTTTTTCAGCAGGACTCGC  
 TTCAGCTTCGCTCGGTATCGGTGATTCATTCTGCTAACAGTAAGCAACCCCGCAGCTAGCCGGTCTCAACGACAGGAGCAGCATGTCGCACCCGTTGGGCGCCCATGCC  
 GCGGATAATGGCTGCTTCTCGCCGAAACGTTGGTGGCGGACCACTGACGAAAGCTTGAAGCGGGCGTGAAGATTCCGAAATACCGCAAGCGACAGGCGCATCTGCTCGCCCTC  
 CAGCGAAAGCGGCTCTCGCCGAAATGACCCAGAGCGCTCGCGCACCTGCTCAGGTTGCATGATAAAGAAGCAGTCATAAGTCCGCGCAGCATAGCTATGCCCGCGCCACC  
 GGAAGGAGCTGACTGGCTTGAAGGCTCAAGGGCATCGCTCAGGATCCCGGCTCAATAGTGAAGCTAATTAATTCGCTTCCGCTCAGCTCCCGCTTCCAGTCCGGGAAA  
 CCTGCTGTCGACGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGGGGTTTTGCTATTGGGCGCAGGGTGGTTTTTCTTTTACCAGTGAGACGGCAACAGCTGATTGCCCTT  
 CACCGCTTGCCCTGAGAGAGTTGCAGCAAGCGGTCACCGTGGTTTCCCGCAGCAGCGCAAAATCTGTTTGTAGTGGTAAACGGCGGGATATAACATGAGCTGCTTCCGATTCG  
 TCGTATCCCACTACCGAGATATCCGCAACCGCGCAGCCGACTCGGTAATGGCGCATTGCGCCAGCGCCATCTGATCGTTGGCAACCGACATCGCAGTGGGAACGATGCCCT  
 CATTACGATTTGCATGGTTTTGTA AAAAACCGGACATGGCACTCCAGTCCGCTTCCCGTTCGCTATCCGCTGAATTTGATGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCGAG  
 ACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTGCTGGTGACCCAATGCGACAGATGCTCCACGCCAGTCCGCTACCTGCTTTCATGGGAGAAAAATAACTGTTG  
 ATGGGTGCTGGTTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCTGCTCATCCAGCGGATAGTTAATGATCAGCCACTGACCGGTT  
 CGCGGAGAAGATTGTGACCCCGCTTTACAGGCTTCGACGCGCTTCTGTTCTACCATCGACACCACCGCTGGCAGCCAGTTGATCGCGCGGAGATTAAATCGCCCGCAAAATTTG  
 CGACGCGCGGTGACGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTTTTGGCCGCGGTTGGGAAATGTAATTCAGCTCCCGCATCGCCGCTTCC  
 ACTTTTTCCCGGTTTTTCGCAAAAACGTTGGCTGGCTGGTTACCACGCGGGAAACGGTCTGATAAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTACATTCA  
 CCACCTGAATTGACTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGGCCATTCGATGGTTCGCGGATCTCGACGCTCTCCCTTATGCGACTCTCGATTAGGAAGC  
 AGCCAGTAGTAGTTGAGGCCCTTGTAGCACCGCCGCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCAAACAGTCCCGCGCCAGCGGGCTGCCACCATACCCAGCCGAAACA  
 AGCCTAGCTAGCGCAAGTGGCGAGGCCATCTCCCATCGGTGATTCGCGGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGATGCGCGGCCAGTCTCGCGCGGTAG  
 AGGATCCAGATCTCGATCCCGGAAATTAATACGACTCACTAAGGGGAATTTGAGCGGATAACAATTTCCCTCTAGAAAATAATTTGTTTAACTTTAAGAAGGAGATATACCATGG  
 TGGGAGTGTACAGGCTATATCGACCGTACAACAGGCGAATGGCATGAAGCAAGCCTTCTATCATCAACAGATGATGGAGATCATGGAGATGAGTCTATCCAGTGGGCAAGCGAG  
 AGACACATTCCTACAGTTAAGCACCAGTCTACGCTGAGGGTACAACATTCCTCGAAGGACCTTATGAGCTGGATGCTTCCAAGGAGAAGACATACCAGAACACAACCTCCCGCGGTTG  
 ACGGCGAAGTTGAGTGGTCAAGCTCGAAGCAAGGAAAGTTGCTATAAAGTTCCAGCTTCAACTCTGTTCTCTCAGACGCTTCTTACGGCGGCTGGACTGAAATGAAGCC  
 TTACGATATTGATAAGGAAACCGCATTTGCTTACTACAACATGAGTTAGGTTCCAGACTTCTGGGGCGACCATCAACAACTATCGCTCAACAACTAAGACTCCTAAGCTCACA  
 ACTGTTGAGTCTGTAATATCCTTGCAGCTCCTGAGGGTGAATCAAGGAGCCTGAGGCTACTTCAAAGATCAAGAAGATCAACCTCAAGGATGCTAAGAACGAGGATACTCTCTATG  
 TTAACCTCGAAGGCGCACCATCAACTAAGACAAACGGTCTCTCGGCTTCATGAAGGGCAGCAGTGGACACAGATCGAGTGGAGCGGTTCTACTGATGCTGACGGCAAGCTCACAGT  
 GGAGATCCCACTTGCTGACGCTGATGAGCGGAACAGTTGAGTCCAGATCTGGCGAGGCTTCAAGGATCTCGATGCAAGGATTACTCAATCGTTGCCATGGCGAGCAGCATCAT  
 CATCATCATCACAGCAGCGCCCTCAGCAAGGGCTGAGG/✂/CCTCAGCTTCCGCTGAGTCCGCTGCACAAGCTTGGCGCCGCTCAGCAGCACCCACCCACCCAC TGAGATCC  
 GGCTGCTAACAAAGCCGAAAGGAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCTCTAAACGGGCTCTGAGGGGTTTTTCTGTAAGGAGGA  
 ACTATATCCGGAT

### pHTP16 sequence landmarks:

- **T7 promoter:** in gray
- **First ATG (methionine):** in yellow
- **CpA gene:** in green
- **His\*Tag coding sequences:** in purple
- **Cloning region:** ✂
- **T7 terminator:** in dark gray
- **Sequencing primers (T7 universal and T7 terminator):** underlined
- **BglII, NcoI & XhoI recognition sites:** in bold

### Sequence added to the final recombinant protein (31.94 kDa):

MGGECHGYIDRTTGEWHEASLP I INKMM EIMDD ESIQWASERH I PTVKHQS YAE GTFLEGP YELDASKEKTY QNTTPGGDGEVEVSWQLEGEVA I KFTGSTP  
 VLCSFSDASYGGWTEMKPYD IDKENG IAYYINMAKVPDLWGGDPTTIAHQAKTPKLTIVESVNI LAPEGE I KEPEATSKIKK INLKDAKNE DTLVNLGEGAP S  
 KTKNGALGFMKGDEWQIEWSGSTADGKLTVE I PLADAVVGGTVEFQ I WAGFKDLVDKYS IVAMGSSHHHHHHSSGPPQQGLR