Isolation of cDNAs that cover the entire coding region of a novel human protein D40

Masato Takimoto, Gang Wei, Pei-zhong Mao, Richard Chikara Koya, Tetsuya Miura and Noboru Kuzumaki
Division of Gene Regulation, Cancer Institute, Hokkaido University School of Medicine, Kita15, Nishi7, Kita-ku, Sapporo, Hokkaido 060-8638, Japan

ABSTRACT
Previously, we have reported a novel protein, D40, that specifically binds to a nuclear factor GCF by two-hybrid system. Northern analysis on tumor cell lines revealed that mRNA of D40 is about 5 kb. To get cDNA clones that cover the total coding region of D40 protein, approximately 5 x 10^5 clones of a human cDNA library were screened, and then several positive clones were obtained. RACE was performed to get the 5' end of the cDNA. DNA sequencing revealed an open reading frame that encodes a protein with 887 amino acids. In vitro transcription and translation analysis of D40 revealed that the molecular weight of the protein is 110 ~ 130 kda.

INTRODUCTION
It has been shown that a transcription factor, GCF(GC Factor), binds to a specific GC-rich sequence in the EGFR promotor and other promoter of growth factors and growth factor receptor genes and negatively regulates the expressions of these gene, such as EGFR, TGF-α, IGF-1R(1-3). Recent results, however, indicates that GCF is a fusion molecule with unrelated cDNA in its 5' end, being probably produced in the process of the construction of cDNA library (4). Furthermore, we revealed the sequence errors in GCF cDNA that produce a frame shift of amino acid sequence of this protein(5).

What is certain is that GCF is a nuclear protein, which was revealed by the study using a specific antibody to the carboxy terminal region of GCF(6). The biological role of GCF has to be re-examined. Identification of a protein that binds specifically to GCF is one of the clues for revealing the biological function of this protein.

We previously revealed the presence of a novel protein that binds specifically to GCF and termed as D40. Simultaneously we isolated a part of cDNA for the protein, by the use of two-hybrid system(7,8). In this paper, we report the isolation of cDNA clones that cover the entire coding region of D40.

MATERIALS and METHODS
Northern analysis. PolyA(+) RNAs were isolated and purified from a human promyelocytic leukemia cell line, HL60, and a T cell leukemia cell line, Jurkat, by a kit (Invitrogen). The poly A RNA were electrophoresed on 1% agarose gel and blotted to a nylon filter. About 2 kbp Bgl II fragment was radiolabeled with α-32P dCTP with a random priming method, and used as a probe for hybridization. After the hybridization, the filter was washed and processed for detection with an image analyzer( BAS2000 FujiFilm).

Screening of cDNA library. Using the Bgl II fragment, HL60 cDNA library made in pcDNAI was screened for clones that cover the full length cDNA of D40. Out of about 5 x 10^5 clones screened, several positive clones were positive and they were isolated, and then DNA sequences were determined.

Rapid Amplification of cDNA Ends (RACE). Total RNA were extracted and isolated from HL60 and Jurkat cell lines with acid guanidinium phenol chloroform(AGPC) method using TRIzol reagent(GIBCO/BRL). RACE method(9) was performed with the protocol and primers provided in 5'3' RACE kit(Boehringer/Roche). For 5' RACE, a D40 specific primer, MT194, was used for
synthesis of the first strand cDNA. The RACE products were analyzed on an agarose gel and then purified from the gel and subcloned in plasmid DNA. DNA sequence were determined.

**DNA sequencing.** Plasmid DNAs were prepared from bacteria with Qiagen tip (QIAGEN). Cycle sequencing reaction was performed with dye terminator method (10), followed by analyses on automatic fluorescent DNA sequencer, ABI PRISM 310.

**Sequence homology search.** Homology search was performed using a BLAST program and data base in National Center for Biotechnology Information (NCBI) at NIH through Internet.

**In vitro transcription and translation.** The cDNA sequence that covers entire coding region of D40 was cloned in pBluescript KS(-) (Stratagene). Using this plasmid as template, in vitro transcription and translation was performed in the presence of 35S-methionine with T7 RNA polymerase and TNT rabbit reticulocyte lysate (Promega).

**RESULTS and DISCUSSION**

In order to isolate cDNAs for D40 that cover the total coding region of this protein, HL60 cDNA library was used for screening. Out of $5 \times 10^5$ yeast clones screened, several positive clones were isolated. Sequence analysis, however, indicated that none of them contains the neither end of the cDNA. To obtain the 5' ends of cDNA, RACE method were applied. Successfully, the 5' end of the cDNA was cloned.

Sequencing the cDNA clones of D40 disclosed a long open reading frame with a capacity of encoding 887 amino acid. There is a in frame stop codon close to the 5' end of the cDNA. Homology search of the amino acid sequence of D40 revealed that here is no protein with significant homology to D40 protein.

In vitro transcription and translation of D40 cDNA, two bands with molecular weight between 110-130 kda were observed in a protein gel. The size of the bands conform to the number of amino acids deduced from cDNA sequence of D40. Presently, it is not clear why two bands were observed for D40. As a fusion protein of D40 with Gal4 DNA-binding domain in its N-terminal end also give two bands, we cannot rule out the possibility that the C-terminal end is missing the translated product of D40 in the lower band.

There are a couple of characteristics in the amino acid sequence of D40. As GCF has a leucine-zipper region in the middle of the protein, it is important to see whether D40 also has leucine-zipper region. However, we were not able to observe leucine-zipper region in the amino acid sequence of D40. As the cDNA clone obtained by two-hybrid system encodes the carboxymal two thirds of D40 protein, the binding domain to GCF must be present in that region.

**ACKNOWLEDGEMENTS**

This work was supported by a Grant in Aid by Ministry of Education, Culture, Science, and Sports in Japan. We thank Dr. S. Kudo for HL60 cDNA library.

**REFERENCE**

Author index

Tanaka, R. 227
Tanaka, T. 215, 217, 259
Tanaka, Yoichiro 221
Tanaka, Yoshiyuki 219, 283
Tanaka, Yuetsu 227
Tao, Z.-F. 249
Tasaka, M. 111
Taya, T. 51
Teramichi, R. 45
Terashima, Y. 35
Terato, H. 83, 291
Tohnai, N. 91, 143
Torigoe, H. 137, 251
Torizawa, T. 37
Tsuge, O. 149
Tsuruoka, M. 59, 239
Uchida, R. 247
Uchida, T. 165
Uchimaru, T. 157, 241
Uda, J. 13
Ueda, Manami 97
Ueda, Masanori 57, 243, 245
Ueda, Mitsunori 165
Ueda, T. 77
Ueki, S. 175
Ueno, A. 271
Ueno, Y. 11, 49, 89, 121
Uesugi, S. 221, 269
Umemura, K. 235
Umeno, D. 55
Umezawa, K. 19
Urata, H. 45
Usui, D. 43
Wada, Takehiko 145
Wada, Takeshi 15, 161
Wake, N. 223
Wang, A.-H.-J. 151, 171
Warashina, M. 219, 283, 285
Washizu, M. 245
Watanabe, A. 193
Watanabe, H. 97
Watanabe, K. 77, 191, 203, 233
Watanabe, M. 101
Watanabe, S. 191, 203
Wataya, Y. 89, 133
Wei, G. 69, 71
Wisniewski, K. 9, 30
Wisniowski, W. 30
Yamada, Masaki 7, 155
Yamada, Masanori 103
Yamada, Mitsuko 97
Yamada, N. 85
Yamada, R. 205
Yamaguchi, T. 205
Yamamoto, H. 103
Yamamoto, K. 93, 95
Yamamoto, N. 227
Yamamoto, R. 269
Yamamoto, S. 215
Yamamura, Shosuke 5, 19
Yamashita, Kamejiro 85
Yamashita, Kenichi 149, 185
Yamashita, Shunsuke 279
Yamato, Y. 23
Yamayoshi, A. 223
Yamazaki, T. 191
Yan, H. 87
Yatomo, C. 25
Yatsunami, R. 73, 75
Yokogawa, T. 133
Yokomatsu, T. 27
Yokotsuka, K. 47
Yokoyama, H. 267
Yokoyama, K. 169
Yokoyama, K.K. 207, 263
Yokoyama, S. 125, 191, 203
Yoneyama, M. 179
Yoshida, J. 71
Yoshida, K. 65, 67
Yoshida, T. 135
Yoshikawa, K. 61
Yoshikawa, Y. 61
Yoshinari, K. 219, 283
Yumoto, T. 45
Zama, M. 81
Zhou, J.-M. 219
PUBLICATIONS IN THE NUCLEIC ACID SYMPOSIUM SERIES

Symposium on RNA Biology III
Proceedings of the 1999 Miami Bio/Technology Winter Symposium
Twentyfifth Symposium on Nucleic Acids Chemistry
Twentyfourth Symposium on Nucleic Acids Chemistry
Symposium on RNA Biology II
Twentythird Symposium on Nucleic Acids Chemistry
Twentysecond Symposium on Nucleic Acids Chemistry
Symposium on RNA Biology I
Proceedings of the 1994 Miami Bio/Technology European Symposium
Twentyfirst Symposium on Nucleic Acids Chemistry
Proceedings of the 1994 Miami Bio/Technology Winter Symposium
Twentieth Symposium on Nucleic Acids Chemistry
Proceedings of the 1993 Miami Bio/Technology Winter Symposium
Nineteenth Symposium on Nucleic Acids Chemistry
Synthetic Oligonucleotides: Problems and Frontiers of Practical Application

ISSN 0261 3166