BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **253,** 16–20 (1998)
ARTICLE NO. RC989725

Polar Amino Acid-Rich Sequences Bind to Polyglutamine Tracts

Ichiro Imafuku,* Masaaki Waragai,* Sosuke Takeuchi,* Ichiro Kanazawa,* Masahiro Kawabata,† M. Maral Mouradian,‡ and Hitoshi Okazawa*,[1]

**Group of Molecular Neurobiology, Department of Neurology, Graduate School of Medicine, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113, Japan; †Department of Biochemistry, The Cancer Institute, Japanese Foundation for Cancer Research, 1-37-1, Kamiikebukuro, Toshima-ku, Tokyo 170, Japan; and ‡Genetic Pharmacology Unit, Experimental Therapeutics Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892*

Received October 23, 1998

Polyglutamine tracts are found in different proteins including transcription factors and cofactors as well as in triplet repeat disease gene products. To characterize the protein motif that binds to the polyglutamine tract, we screened a human embryonic brain cDNA library with the polyglutamine tract of Brn-2 as bait using the yeast two-hybrid method. All six isolated clones encoding polyglutamine tract binding proteins were rich in polar amino acids. Three of these clones could form polar helical structures. These observations suggest that polar amino acid-rich sequences are essential for binding to the polyglutamine tract.

Polyglutamine tracts are present in various proteins including triplet repeat disease gene products, transcription factors, and transcriptional co-factors. This sequence module has been suggested to form a polar zipper structure (1) and to function as a transactivation domain in Oct-2 (2). Expanded polyglutamine tracts in triplet repeat disease proteins self-aggregate to form nuclear inclusion bodies toxic to neurons (3–5). The self aggregation step is thought to be modulated by transglutaminase (6, 7). These observations implicate the polyglutamine tract in protein–protein interactions under physiologic and pathologic conditions.

Several proteins interacting with triplet repeat disease gene products have been isolated (9–13) but their direct binding to the polyglutamine tract itself has not been demonstrated. Some of these interacting proteins have the WW domain which has been suggested to bind to the proline-rich sequence in the two-hybrid screening bait (12, 13). With the exception of the pathological interaction among expanded polyglutamine tracts themselves, the molecular motif necessary for binding to the polyglutamine tract remains unknown. To address this question, we performed yeast two-hybrid cDNA library screening using a bait limited specifically to the polyglutamine tract and thereby isolated clones which bind exclusively to this sequence. Interestingly, all these clones encoded polar amino acid-rich sequences, suggesting that the latter is an essential structural feature for interaction with the polyglutamine tract.

MATERIALS AND METHODS

Yeast two-hybrid interaction. Two-hybrid cDNA cloning was performed according to the method of Gyuris *et al.* (14). To construct the bait plasmid, a DNA fragment corresponding to the polyglutamine tract of Brn-2 was amplified by PCR from pCMVBrn2 (15) and subcloned between the *Eco*RI and *Xho*I sites of pEG202 in the correct frame. The resultant fragment encoded amino acids 122–154 of the murine Brn-2, which in addition to 26 glutamines, includes Gly-Ala-Leu and Arg-Pro-Pro at the 5′ and 3′ ends, respectively, and a His at the 4th position in the polyglutamine tract (16). After confirming that the bait plasmid itself could not activate transcription from the GAL1–GAL10 promoter in the reporter plasmid pSH18-34, we screened a human embryonic brain cDNA library constructed in pJG4-5. Approximately 2×10^6 colonies were screened on plates with synthetic medium lacking histidine, tryptophan and uracil. Positive clones were checked in four types of second plates as described (14). pJG4-5 plasmids were recovered from the final positive yeast clones and their inserts were sequenced using an automated sequencer. To obtain the entire sequence of inserts, they were subcloned in pBS-SK (Stratagene) and the Kilo-Deletion Kit for DNA Sequencing (Takara, Japan) was used.

[1] To whom correspondence should be addressd. Fax: 0081-3-5800-6548. E-mail: okazawa-tky@umin.u-tokyo.ac.jp.

0006-291X/98 $25.00

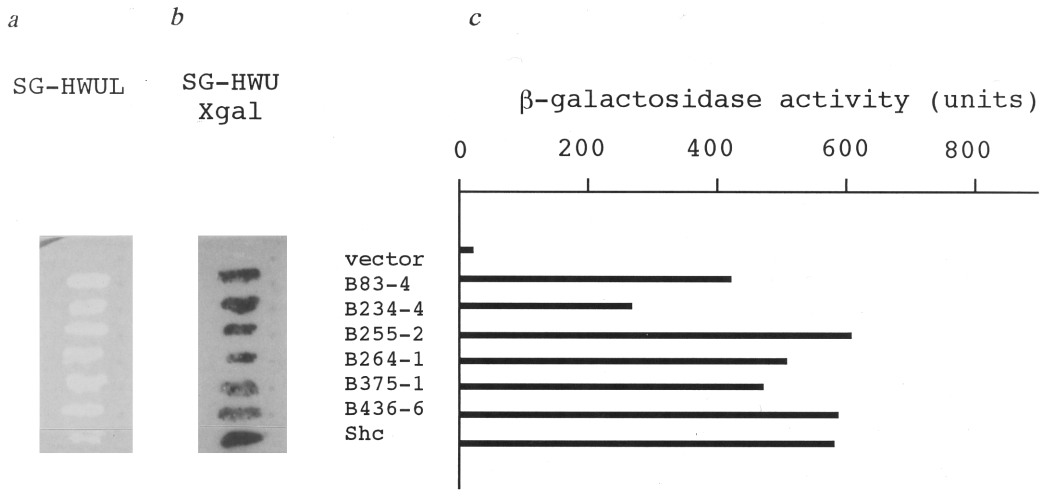


FIG. 1. Two hybrid interaction between the isolated clones and the polyglutamine tract of Brn-2. Interaction was assayed by galactose induction in Leucine (−) or X-gal plates.

RESULTS AND DISCUSSION

We had previously demonstrated that different members of the POU domain transcription factor family have different transactivation potentials and suspected the involvement of various cofactors binding to POU proteins (17, 18). In the present investigation, we hypothesized that the polyglutamine tract in POU proteins might be the interacting site for transcriptional cofactors, since this motif could form a polar zipper structure (1) exposed on the exterior of the POU molecule. To test this hypothesis, two-hybrid screening of a human embryonic brain cDNA library (a kind gift from Dr. Brent) was performed using a bait which contained exclusively the polyglutamine tract of Brn-2 (16).

Among the positive cDNA clones obtained from the second screening, those in the wrong reading frame, of which some encoded known genes in different frames, and those encoding ribosomal proteins were excluded. Six clones which had proper reading frames and showed strong interaction to the bait were judged to be specific (Fig. 1). One of these clones (B264-1) was identical to transitional endoplasmic reticulum ATPase (TERA) which participates in ATP-dependent protein transport from the endoplasmic reticulum to the Golgi apparatus (19, 20). We designated the other clones as poly-Q binding proteins (PQBP) 1 through 5 (Fig. 2). PQBP-1 sequence has been reported in the genome database although its function is not explored. Our studies of PQBP-1 including its expression and functions will be reported elsewhere (Waragai *et al.*, submitted). The remaining four clones had deduced amino acid sequences not homologous to any known sequence as determined by BLAST search (Fig. 2). Some of the isolated clones (PQBP-1, -3, and TERA) are expressed in the human brain by northern blot analysis (Fig. 3). Notably, triplet repeat disease proteins were not isolated in this screen as binding partners to the polyglutamine tract.

Comparison of the amino acid sequences of isolated clones revealed that all are rich in polar amino acids such as arginine, lysine, histidine, aspartic acid and glutamic acid. The basic or acidic side

clone No.	cDNA size	BLAST homology search	motif/character	designation
B83-4	1.0Kb	unknown	Arg, Asp-rich	PQBP-1
B234-4	1.8Kb	unknown	Arg, Asp, Glu-rich	PQBP-2
B255-2	1.0Kb	unknown	Arg-rich	PQBP-3
B264-1	1.0Kb	known	Ser, Asp-rich	TERA
B375-1	0.6Kb	unknown	Arg-rich	PQBP-4
B436-6	1.3Kb	unknown	Arg, Ser, Lys-rich	PQBP-5

FIG. 2. List of cDNA clones isolated by two-hybrid screening. cDNA insert size, BLAST homology search result and amino acid sequence characteristics are indicated. TERA is transitional endoplasmic reticulum ATPase.

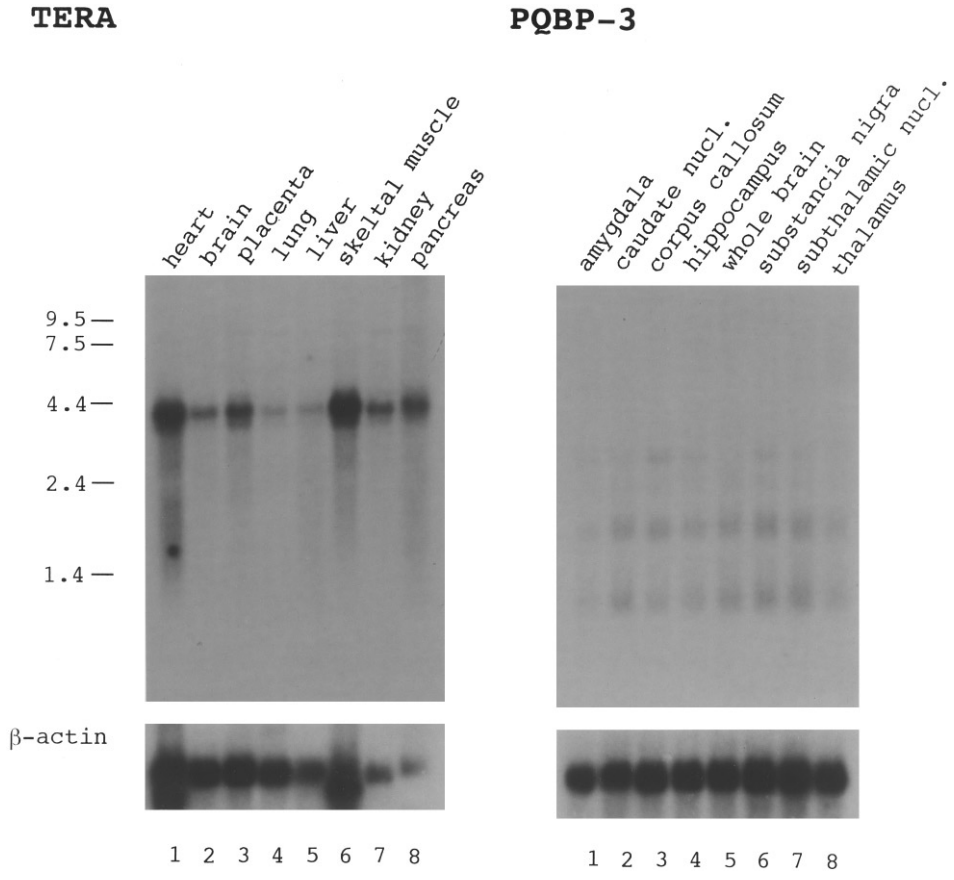


FIG. 3. Expression of TERA and PQBP-3 in the human brain.

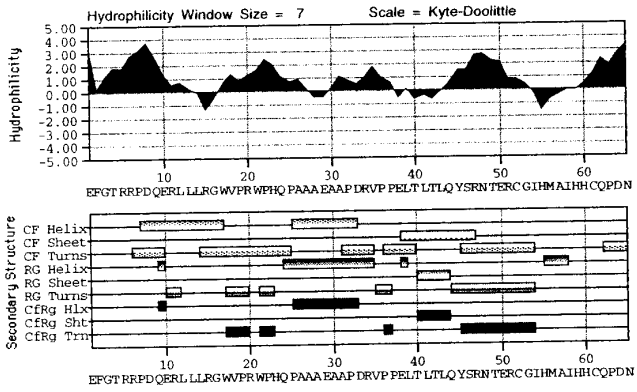
chains in these amino acids are highly polar and are expected to position on the exterior of the protein molecule. Consistent with this notion, hydrophilicity analysis using MacVector II indicated that these sequences are highly hydrophilic (Fig. 3). These polar amino acids form helical structures in PQBP-1, PQBP-2 and PQBP-5. Furthermore, we have found that this structure in PQBP-1 is essential for binding to the polyglutamine tract based on deletion analysis (Waragai *et al.*, submitted). Although PQBP-3, PQBP-4, and TERA are not predicted to form helices, they too are very hydrophilic (Fig. 4).

Polar amino acid-rich sequences in general and polar helical structures in particular appear to be capable of interacting with the polyglutamine tract. This is a plausible conclusion since the polyglutamine tract is predicted to form a polar zipper (1) which would be sticky to polar structures via hydrogen bonds. Absence of polyglutamine tract containing proteins among our isolated clones could be due to the fact that the polarity of these tracts are not as high as the polar-rich sequences of our PQBPs. This is because glutamine with an uncharged polar side chain is less polar than amino acids with acidic or basic side chains. Although higher polarity is acquired by CAG expansions, the normal length of the polyglutamine tract might be insufficient to bind to its protein partners in the two hybrid assay. Polar amino acid rich sequences could probably bind to the polyglutamine tract under physiologic conditions, but expansion of the triplet repeat and some additional factors might be necessary for self-interaction among polyglutamine tracts. The latter could have pathophysiological implications in the genesis of protein aggregates in triplet repeat disorders.

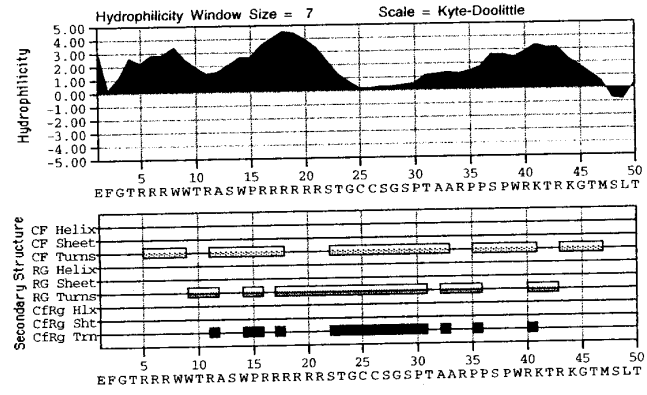
ACKNOWLEDGMENTS

We thank Dr. Roger Brent (Department of Genetics, Harvard University) for the two-hybrid system and Dr. Hiroshi Hamada (Institute for Molecular Biology, Osaka University) for Brn-2 cDNA. This work was supported by Grants from the Japanese Ministry of Education, Culture, Science, and Sports (07558233, 10670574) to H.O.

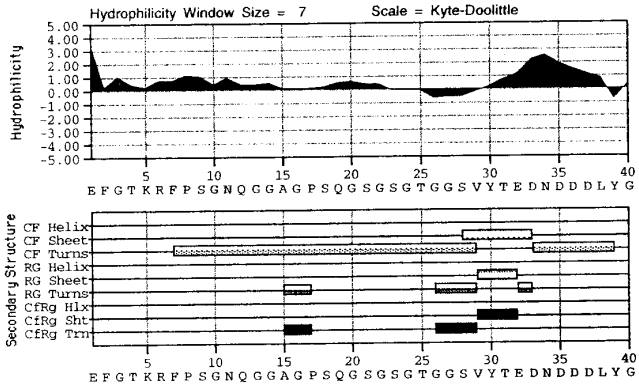
PQBP-2



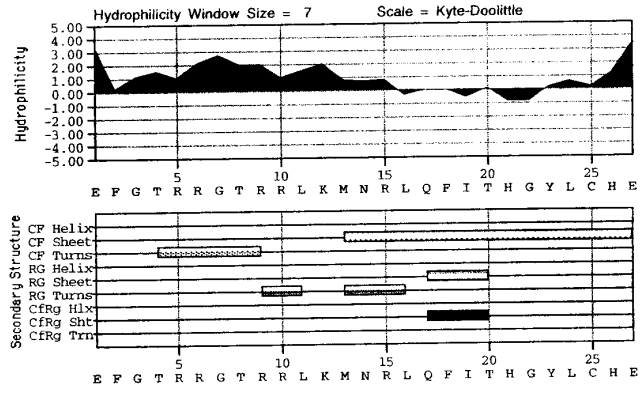
PQBP-3



TERA



PQBP-4



PQBP-5

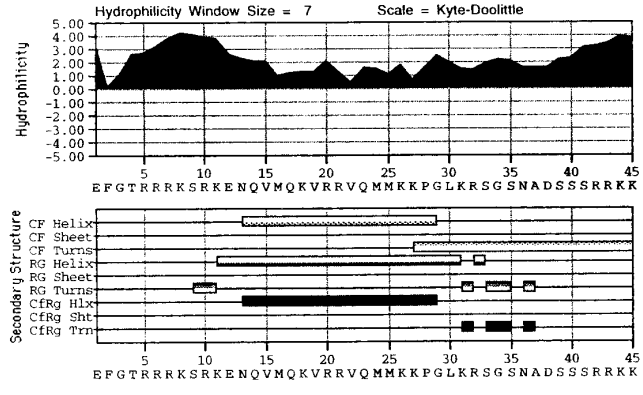


FIG. 4. Hydrophilicity and secondary structure analyses of PQBPs and TERA.

REFERENCES

1. Perutz, M. F., Johnson, T., Suzuki, M., and Finch, J. T. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5355–5358.
2. Tanaka, M., and Herr, W. (1990) *Cell* **60,** 375–386.
3. Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenbank, R., Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997) *Cell* **90,** 549–558.
4. Paulson, H. L., Perez, M. K., Trottier, Y., Trojanowski, J. Q., Subramony, S. H., Das, S. S., Vig, P., Mandel, J. L., Fischbeck, K. H., and Pittman, R. N. (1997) *Neuron* **19,** 333–344.
5. DiFiglia, M., Sapp, E., Chase, K. O., Davies, S. W., Bates, G. P., Vonsattel, J. P., and Aronin, N. (1997) *Science* **277,** 1990–1993.
6. Kahlem, P., Green, H., and Djian, P. (1998) *Mol Cell.* **1,** 595–601.
7. Igarashi, S., Koide, R., Shimohata, T., Yamada, M., Hayashi, Y., Takano, H., Date, H., Oyake, M., Sato, T., Sato, A., Egawa, S., Ikeuchi, T., Tanaka, H., Nakano, R., Tanaka, K., Hozumi, I., Inuzuka, T., Takahashi, H., and Tsuji, S. (1998) *Nature Genet.* **18,** 111–117.
8. Li, X. J., Li, S. H., Sharp, A. H., Nucifora, F. C., Jr., Schilling, G., Lanahan, A., Worley, P., Snyder, S. H., and Ross, C. A. (1995) *Nature* **378,** 398–402.
9. Kalchman, M. A., Koide, H. B., McCutcheon, K., Graham, R. K., Nichol, K., Nishiyama, K., Kazemi-Esfarjani, P., Lynn, F. C., Wellington, C., Metzler, M., Goldberg, Y. P., Kanazawa, I., Gietz, R. D., and Hayden, M. R. (1997) *Nature Genet.* **16,** 44–53.
10. Burke, J. R., Enghild, J. J., Martin, M. E., Jou, Y. S., Myers, R. M., Roses, A. D., Vance, J. M., and Strittmatter, W. J. (1996) *Nature Med.* **2,** 347–350.
11. Matilla, A., Koshy, B. T., Cummings, C. J., Isobe, T., Orr, H. T., and Zoghbi, H. Y. (1997) *Nature* **389,** 974–978.
12. Wood, J. D., Yuan, J., Margolis, R. L., Colomer, V., Duan, K., Kushi, J., Kaminsky, Z., Kleiderlein, J. J., Sharp, A. H., and Ross, C. A. (1998) *Mol. Cell Neurosci.* **11,** 149–160.
13. Farber, P. W., Barnes, G. T.,Srinidhi, J., Chen, J., Gusella, J. F., and MacDonald, M. E. (1998) *Hum. Mol. Genet.* **7,** 1463–1474.
14. Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. (1993) *Cell* **75,** 791–803.
15. Fujii, H., and Hamada, H. (1993) *Neuron* **11,** 1197–1206.
16. Hara, Y., Rovescalli, A. C., Kim, Y., and Nirenberg, M. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 3280–3284.
17. Imafuku, I., Kamei, M., Kanazawa, I., Mouradian, M. M., and Okazawa, H. (1996) *Biochem. Biophys. Res. Commun.* **222,** 736–741.
18. Okazawa, H., Imafuku, I., Minowa, M. T., Kanazawa, I., Hamada, H., and Mouradian, M. M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11933–11938.
19. Egerton, M., Ashe, O. R., Chen, D., Druker, B. J., Burgess, W. H., and Samelson, L. E. (1992) *EMBO J.* **11,** 3533–3540.
20. Zhang, L., Ashendel, C. L., Becker, G. W., and Morre, D. J. (1994) *J. Cell Biol.* **127,** 1871–1883.