Vol. 150, No. 1, 1988 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1988 Pages 463-468

# SIMILARITY BETWEEN THE ESTROGEN RECEPTOR AND THE DNA-BINDING DOMAIN OF THE TETRACYCLINE REPRESSOR

Michael E. Baker

Department of Medicine, M-023
University of California, San Diego
La Jolla, California 92093

Received November 30, 1987

**Summary.** The amino-terminus region of tetracycline (*tet*) repressors contains a helix-turn-helix structure that binds to DNA. A computer-based comparison of a residues 18-61 of the *tet* repressor, which contains this DNA-binding domain, with a residues 304-347 of the human estrogen receptor (ER) yields a score that is 6.7 standard deviations higher than that obtained with 2,500 comparisons of randomized sequences of these segments. The probability of getting this score by chance is less than $10^{-11}$. This part of the ER could be important in nuclear actions of ER in target tissues. 

Steroids act through a specific receptor protein in a target tissue (1,2). The binding of the steroid to the receptor effects a conformational change in the receptor, which increases its affinity for specific nuclear binding site(s) (3-5). The binding of the hormone-receptor complex to these nuclear sites induces the synthesis of various gene products. The details of how steroids bind to receptors and increase the affinity of the receptor for nuclear sites, and how the steroid-receptor complex induces gene transcription are still poorly understood. The cloning and sequencing of genes for glucocorticoid receptor (GR) (6,7), estrogen receptor (ER) (8-10), progesterone receptor (PR) (11), mineralocorticoid receptor (MR) (12), and thyroid hormone receptor (TR) (13,14) opens the door for understanding the details of the molecular biology of steroid hormone action.

Examination of the sequences of these receptors reveals a highly conserved ~65 residue sequence near their middle, which contains several cysteines and basic residues. It was proposed that this segment, corresponding to residues 185-250 of the human ER, functioned as a DNA-binding domain (8-10) because it contains cysteines that could form finger-like structures analogous to those in proteins from other organisms that use a cysteine-histidine-zinc finger-like motif to bind DNA (15-19). It should be noted that this hypothesis was based on a chemical analogy because there is no statistically significant similarity between the sequence of the finger-like structure on DNA-binding proteins and steroid receptors. Recent site-specific mutagenesis studies support the hypothesis that residues 185-250 of the ER bind DNA (20,21).

We were intrigued by the absence of the DNA-binding structure found in prokaryote repressors and activators (22-24) on steroid receptors. To investigate this more thoroughly, regions on the ER, GR, PR,

0006-291X/88 $1.50

MR, and TR were examined for similarity to prokaryote repressors. Here we present evidence that residues 304-347 of the human ER are similar to a segment near the amino-terminus of the tetracycline repressor (*tet* repressor), a DNA-binding protein that regulates transcription of plasmid-encoded tetracycline resistance genes in *Escherichia coli* and other enteric bacteria. This is the segment of *tet* repressor that is thought to be involved in binding DNA (25-29). If this *tet* repressor-like sequence on the ER interacts with DNA or its associated proteins (3-5), then the action of the ER is more complex than previously thought.

**Methods.** The RELATE and ALIGN computer programs developed at the National Biomedical Research Foundation (30,31) were used to compare the ER, GR, PR, MR, and TR with sequences of prokaryote repressor/activator proteins in the Dayhoff database, which include catabolite gene activator protein, and λ, λ Cro, Lac, Trp, Gal, LexA, Transposase, Biotin-operon, and *tet* repressors (23,24). No similarities were found between the GR, PR, MR, or TR and the *tet* repressor or the other repressors in the database.

**Results.** The RELATE analysis revealed that a segment near the amino terminus of *tet* repressors (25,26,28) was similar to residues 304-347 of the ER. The 4 *tet* repressors, for which sequence data are available, are all similar to this part of the ER. As an example, the alignment between residues 18-61 of *tet* repressor of plasmid pSC101 (25) and the ER is shown in Figure 1. The alignment reveals that, of the 44 residues in these segments, there are 9 identities (20%), 7 conservative replacements (16%), and 18 residues that differ in the genetic code by only one base (41%).

A quantitative measure of this similarity was obtained with the ALIGN computer analysis, which yielded a comparison score that is 6.7 standard deviations higher than that of 2,500 comparisons of randomized sequences of these segments. The probability of getting this score by chance is less than $10^{-11}$. To get a perspective on this score, pSC101 *tet* repressor was compared with other prokaryote DNA-binding proteins, using the RELATE program. The score with the other *tet* repressors was about 40 standard deviations above that of comparisons of randomized sequences. In contrast, the RELATE comparison of pSC101 *tet* repressors with other DNA-binding proteins (23,24) did not yield a score above 3 standard deviation units, which indicates substantial divergence of the primary amino acid sequence of the *tet* repressor from repressors in other organisms. Sauer et al. (23) also found that repressor/activators from different organisms have diverged considerably. This makes the similarity between *tet* repressor and the ER even more intriguing.

**Discussion.** Residues 26-47 of all *tet* repressors that have been sequenced are highly conserved and are thought to constitute their DNA-binding site (29). Mutations in this region of the Tn*10* repressor abolish DNA binding without significantly altering tetracycline binding or subunit interaction (29). Residues 310-332 of the ER correspond to this region on the *tet* repressor (Figure 1), which suggests that this part of the ER may be involved in binding to nuclear acceptor sites in target tissues.

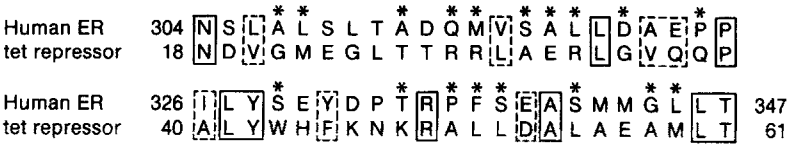


Figure 1. Alignment of human estrogen receptor and pSC101 *tet* repressor.

Alignment of residues 304-347 of human ER with residues 18-61 of pSC101 *tet* repressor. Solid boxes show identities and the dotted boxes show conservative replacements according to the scheme: (P,G), (M,C), (Y,W,F,H), (L,V,I,A), (K,R), (E,Q,N,D), and (S,T). Other amino acid differences which can be explained by a single nucleotide change in the codon are indicated by an asterisk. The ALIGN comparison of these segments yields a score that is 6.7 standard deviations higher that that of 2,500 comparisons of randomized sequences of these segments. The probability of getting this score by chance is $10^{-11}$.

Prokaryote DNA-binding proteins contain a characteristic helix-turn-helix structure at their DNA-binding site (22-24). With the caveat that computer-based predictions of secondary structure are still being perfected, an analysis of the secondary structure of the ER by Krust et al. (10) indicates the presence of a helix-turn-helix structure in residues 314-343 of the ER. This analysis suggests that residues 314-323 and 338-343 are alpha helicies, and residues 330-337 are a turn. While the secondary structure prediction for residues 314-343 of the ER is consistent with that needed for a DNA-binding structure, it should be noted that compared to the prokaryote DNA-binding domains, the region on the ER between the proposed helicies is longer and does not contain a highly conserved glycine, which is replaced with an aspartic acid residue. This difference is not unexpected considering that chromatin acceptor sites for the ER are more structurally complex than the DNA-binding sites in prokaryotes (3-5).

Residues 185-250 of the ER, which comprise the finger-like domain, are the primary determinant of DNA binding *in vitro* (21). A similar result has been found for the finger-like domain of the GR (32). From these *in vitro* studies it would appear that if residues 304-343 of the ER are involved in its binding to nuclear acceptor sites, their role is secondary to that of the finger-like domain. However, these studies were done with *in vitro* systems. For more diverse and complex *in vivo* systems (3-5), parts other than the finger-like domain of the ER may have a role in binding to nuclear acceptor sites.

In fact, there is evidence from research with the rat GR, which is likely to have a similar mechanism of action as the ER, supporting this possibility. Two nuclear localization sites have been found on the GR (33). One of these sites in the rat GR (residues 497-524) contains a highly basic stretch of amino acids, which are similar to the known nuclear localization site of SV40 large T antigen (34,35). At a corresponding position of the ER, there is a highly basic stretch of amino acids, which are part of a domain (residues 255-281), that has been shown to be similar to sturgeon protamine, a DNA-binding protein (36).

The other site is in the GR's steroid binding domain, although the exact location has not yet been determined. Residues 304-347 of the ER overlap the estrogen binding site. The possibility of two nuclear localization sites on the ER was first suggested by Wolff et al. (37) based on their comparison of steroid receptor sequences with the known nuclear localization site for SV40 large T-antigen. One of these sites (residues 263-269) is part of the protamine-like segment of the ER mentioned previously; the other, residues 298-304, is adjacent to the *tet* repressor-like part of the ER. Both segments contain basic residues, which could interact with negatively charged sites in the nucleus. The observations concerning nuclear localization sites on the GR and the hypothesis of Wolff et al. about the nuclear localizations sites on ER support the notion that the *tet* repressor-like segment of the ER has a role in nuclear actions of the ER.

Site specific mutagenesis is one way to study the biological role of residues 304-347 of the ER. This could clarify whether this segment has a role in binding to nuclear acceptor sites or if it has some other function. Because the alpha helical structure is central to binding to DNA of prokaryote repressor/activators, Eco RI endonuclease (38), and possibly the finger-like structures (39), disruption of the helicies in residues 304-347 of the ER, by inserting a helix-breaking proline residue, would be one way to begin to determine whether residues 304-347 of the ER are important in its binding to chromatin acceptor sites.

This *tet* repressor-like part of the ER may have other roles in the *in vivo* actions of estrogen. Residues 304-347 of the ER are close to and may partially overlap the steroid binding site (10,20,36). If these residues interact with or are close to nuclear acceptor sites, then it is possible that either the steroid and/or its binding site could interact with DNA (40) or its associated proteins. The first possibility, that receptor-bound steroid intercalated in DNA could have biological importance was noted about 25 years ago by Huggins and Yang (41). Support for their hypothesis came from Weeks et al.'s (42) x-ray crystallographic studies of deoxycorticosterone-adenine crystals, which revealed steroid intercalation between adenine stacks, as well as the presence of hydrogen bonding between deoxycorticosterone and adenine. Of course, steroid intercalation in DNA provides a mechanism for biological specificity.

The second possibility, that the steroid binding site could interact with DNA or its associated proteins, is interesting because of the presence of a nucleophilic site on the ER, which has chemical properties like the those at the catalytic site of proteases, at or close to the steroid binding site (43-46) and because proteins can interact with phosphodiesters (47) and proteins (48) via a nucleophilic site. This interaction could involve catalysis, or it could involve a binding action as has been found in haptoglobin (a serine protease homologue without catalytic activity) and the estero-peptidases that bind to nerve growth factor and epidermal growth factor (49,50).

The computer analyses can not distinguish between the similarity of part of the ER and *tet* repressor being a result of convergent or divergent evolution. Still, it is interesting to note that tetracycline contains some structural similarities to estrogen, including a phenolic ring analogous to estrogen's A ring. Moreover, tetracycline has even stronger resemblence to zearalenone, a fungal compound that has estrogenic activity (51). Both estrogen (52) and zearalenone are found in primitive organisms, and it is possible that they and tetracycline compete for similar binding sites in certain *in vivo* systems.

**Acknowledgements.** I thank Kevin Bertrand, Darrell Fanestil, and Greg Sutcliffe for discussion and Marvin F. Weissberg for encouragement and support.

**References**

1. Yamamoto, K.R. (1985) Annu. Rev. Genet. 19, 209-252.
2. Katzenellenbogen, B.S. (1980) Annu. Rev. Physiol. 42, 17-35.
3. Spelsberg, T.C. (1982) In *Biochemical Actions of Hormones*, Vol. 9 (Litwack, G., Ed.) pp. 141-204, Academic Press, New York.
4. Spelsberg, T.C. (1983) Biochemistry 22, 13-21.
5. Ruh, T.S. and Baudendistel, L.J. (1978) Endocrinol. 102, 1838-1846.
6. Hollenberg, S.M., Weinberger, C., Ong, E.S., Cerelli, G., Oro, A., Lebo, R., Thompson, E.B., Rosenfeld, M.G., and Evans, R.M. (1985) Nature 318, 635-641.
7. Weinberger, C., Hollenberg, S.M., Rosenfeld, M.G., and Evans, R.M. (1985) Nature 318, 670-672.
8. Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J-M., Argos, P., and Chambon, P. (1986) Nature 320, 134-139.
9. Greene, G.L., Gilna, P., Waterfield, M., Baker, A., Hort, Y., and Shine, J. (1986) Science 231, 1150-1154.
10. Krust, A., Green, S., Argos, P., Kumar, V., Philippe, W., Bornert, J-M., and Chambon, P. (1986) EMBO J., 5, 891-897.
11. Loosfelt, H. Atger, M., Misrahi, M., Guiochon-Mantel, A., Meriel, C., Logeat, F., Benarous, R., and Milgrom, E. (1986) Proc. Nat. Acad. Sci. USA 83, 9045-9049.
12. Arriza, J.L., Weinberger, C., Cerelli, G., Glaser, T.M., and Handelin, B.L., Housman, D.E., and Evans, R.M. (1987) Science 237, 268-275.
13. Sap, J., Munoz, A., Damm, K., Goldberg, Y, Ghysdael, J., Leutz, A., Beug, H., and Vennstrom, B. (1986) Nature 324, 635-640.
14. Weinberger, C., Thompson, C.C., Ong, E.S., Lebo, R., Gruol, D.J., and Evans, R.M. (1986) Nature 324, 641-646.
15. Miller, J., McLachlan, A.D., and Klug, A. (1985) EMBO J. 4, 1609-1614.
16. Berg, J.M. (1986) Nature 319, 264-265.
17. Rosenberg, U.B., Schröder, C., Preiss, A., Kienlin, A., Côté, S, Riede, I., and Jäckle, H. (1986) Nature 319, 336-339.
18. Hartshorne, T.A., Blumberg, H., and Young, E.T. (1986) Nature 320, 283-287.
19. Harrison, S. (1986) Nature 322, 597-598.
20. Kumar, V., Green, S, Staub, A., and Chambon, P. (1986) EMBO J. 5, 2231-2236.
21. Green, S. and Chambon, P. (1987) Nature 325, 75-78.
22. Ohlendorf, D.H., Anderson, W.F., Fisher, R.G., Takeda, Y., and Matthews, B.W. (1982) Nature 298, 718-723.
23. Sauer, R.T., Yocum, R.R., Doolittle, R.F., Lewis, M, and Pabo, C.O. (1982) Nature 298, 447-451.
24. Pabo, C.O. and Sauer, R.T. (1984) Ann. Rev. Biochem. 53, 293-321.
25. Waters, S.H., Rogowsky, P., Grinsted, J., Attenbuchner, J., and Schmitt, R. (1983) Nucleic Acids Res. 11, 6089-6105.
26. Postle, K., Nguyen, T.T., and Bertrand, K.P. (1984) Nucleic Acids Res. 12, 4849-4863.
27. Brow, M.A.D., Pesin, R., and Sutcliffe, J.G. (1985) Molec. Biol. and Evol. 2, 1-12.
28. Unger, B., Klock, G., and Hillen, W. (1984) Nucleic Acids Res. 12, 7693-7703.
29. Isackson, P.J. and Bertrand, K.P. (1985) Proc. Nat. Acad. Sci. USA 82, 6226-6230.

30. Barker, W.C. and Dayhoff, M.O. (1982) Proc. Nat. Acad. Sci. 79, 2836-2839.
31. Dayhoff, M.O., Barker, W.C., and Hunt, L.T. (1983) Meth. in Enzymol. 91, 524-545.
32. Godowski, P.J., Rusconi, S., Miesfeld, R., and Yamamoto, K.R. (1987) Nature 325, 365-368.
33. Picard, D. and Yamamoto, K.R. (1987) EMBO J. 6, 3333-3340.
34. Kalderon, D., Richardson, W.D., Markham, A.F., and Smith, A.E. (1984) Nature 311, 33-38.
35. Lanford, R.E., Kanda, P., and Kennedy, R.C. (1986) Cell 46, 575-582.
36. Baker, M.E. (1986) Biochem. Biophys. Res. Commun. 139, 281-286.
37. Wolff, B., Dickson, R.B., and Hanover, J.A. (1987) Trends Pharmacol. Sci. 8, 119-121.
38. McClarin, J., Frederick, C.A., Wang, B-C., Greene, R., Boyer, H.W., Grable, J., and Rosenberg, J.M. (1986) Science 234, 1526-1541.
39. Brown, R.S. and Argos, P. (1986) Nature 324, 215.
40. Sluyser, M. (1983) Trends in Biochem. Sci. 8, 236-238.
41. Huggins, C. and Yang, N.C. (1962) Science 137, 257-262.
42. Weeks, C.M., Rohrer, D.C., and Duax, W.L. (1975) Science 190, 1096-1097.
43. Baker, M.E., Vaughn, D.A., and Fanestil, D.D. (1978) J. Supramol. Struct. 9, 421-426.
44. Lukola, A. and Punnonen, R. (1982) Biochem. Biophys. Res. Commun. 108, 822-827.
45. Baker, M.E., Sklar, D.H., Terry, L.S., and Hedges, M.R. (1985) Biochem. Internat. 11, 233-238.
46. Puca, G.A., Abbondanza, C., Nigro, V. Armetta, I., Medici, N., and Molinari, A.M. (1986) Proc. Nat. Acad. Sci. 83, 5367-5371.
47. Suck, D. and Oefner, C. (1986) Nature 321, 620-625.
48. Neurath, H. (1984) Science 224, 350-357.
49. Baker, M.E. (1987) In *Biological Activities of Alpha-Fetoprotein* Vol. 1 (Mizejewski, G.J. and Jacobson, H.I., Ed.) pp. 35-50, CRC Press, Boca Raton, Florida.
50. Moore, J.B. Jr., Mobley, W.C., and Shooter, E.M. (1974) Biochemistry 13, 833-839.
51. Katzenellenbogen, B.S., Katzenellenbogen, J.A., and Mordecai, D. (1979) Endocrinol. 105, 33-40.
52. Feldman, D., Tökés, L., Stathis, P.A., Miller, S.C., Kurz, W., and Harvey, D. (1984) Proc Nat. Acad. Sci. USA 81, 4722-4726.