ANTISENSE HOMOLOGY BOXES IN PROTEINS

Lajos BARANYI [1] William CAMPBELL [1] Hidechika OKADA [2]
Compuserve:100210.1512 Compuserve:100210.1512 Compuserve:100210.1512

[1] Choju Medical Institute, Fukushimura Hospital
19-14 Aza Yamanaka. Toyohashi. 441 Japan

[2] Department of Molecular Biology Nagoya City University School of Medicine,
Nagoya 467, JAPAN

Abstract

Amphiphilic peptides approximately fifteen amino acids in length and their corresponding antisense peptides exist within protein molecules. These regions (termed antisense homology boxes) are separated by approximately fifty amino acids. Since many sense-antisense peptide pairs have been reported to recognize and bind to each other, antisense homology boxes may be involved in folding, chaperoning and oligomer formation of proteins. The finding that cc 70 per cent of the antisense homology box derived peptides from C5a receptor and human Endothelin A receptor are specific inhibitors or agonist to their corresponding proteins indicate that these regions can have significant role in proteins.

1 Introduction

With the advent of large scale genetic sequencing the amount of availabl e data on amino acid sequences of proteins accumulates exponentially. However the understanding and use of these molecules is hampered by our inadequate knowledge structure-function relationship. Here we report a new motif in proteins with its roots in the genetic code itself. It has already been known for several decades that there is a tendency i n the genetic code for codons of hydrophilic amino acids to be complemented on the complementary DNA strand by codons for hydrophobic amino acids and vice versa. Codons for slightly hydrophilic amino acids are similarly complemented. Since the hydrophobicity of the amino acids is mainly determined by the central nuclei acid, this tendency is valid in the case of antisense codons. Blalock and Smith found significant correlation (r= 0.74) between the average hydropathic coefficients of amino acids encoded by opposing DNA strands. When synthesized, a substantial number of the sense-antisense peptides recognized and bound to each other and furthermore, of reported interacting complementary peptide pairs is steadily increasing. Based on these correlations, a molecular recognition theory (MRT) was proposed by Blalock et al. [2,3]. We extended the principles of the MRT taking into consideration the degeneracy of the genetic code that allows for complementarity within proteins to occur without excess complementarity on the mRNA level [1]. We also present a n example of how to utilize the new concept in order to predict possible "antisense based" intramolecular recognition sites in proteins and their use in the design of biologically active peptides.

2 RESULTS

Sense-antisense relationship within proteins. Using our software "ANTIS" on IBM personal computers, we determined the number and distribution of intramolecular regions having sense-antisense relationships in randomly selected sets of proteins. To provide reference data, the search was repeated using sets of randomly selected amino acids. For each of the proteins during the "random complementary search", the antisense amino acid set was replaced with 3 different, randomly selected amino acid sets. A much higher number of antisense frames was found when the real antisense amino acid set was used than when any of the random amino acid sets were used (total: 787 versus 112, 180 and 270 using random sets 1, 2 and 3, respectively) or 4.2 times more on average (p¡0.001 in each case).

[1]Barany Lajos. Choju Medical Institue. Fukushimura Hospital. 19-14 Aza Yamanaka. Toyohashi. 441 Japan
[2]William Campbell. Choju Medical Institue. Fukushimura Hospital. 19-14 Aza Yamanaka. Toyohashi. 441 Japan

The distance between antisense homology boxes. The expected number of antisense frames can be calculated. The number of antisense boxes found using random amino acid sets correlated well with calculated values. The ratio of the found to the expected number of homology boxes was close to 1. However, when the real antisense amino acid set was used for searching, the number of antisense homology boxes was higher than expected. Interaction between complementary antisense homology boxes may be envisioned as a dynamic docking mechanism with variable binding strength since recently, a novel type of secondary structure was proposed which is formed when complementary peptides interact. Hydrophilic residues of both chains are oriented toward the aqueous solvent, while the hydrophobic ones form the interphase between the two chains. Tight packing is possible because whenever a hydrophilic residue turns toward the aqueous phase, a space appears in the carbon backbone that can accommodate a hydrophobic residue from the opposing chain. This model predicts that complementary antisense homology bo xes can interact strongly when the protein structure is distorted and more weakly in stabilized, closely packed structural elements. The association of antisense homology boxes with amphiphilic regions containing reverse turns implies that most of them reside at positions where the automatic folding units, i.e., helices and beta sheets end or start, therefore, they may serve primarily as "stabilizers" during the folding and unfolding of proteins. Zull and Smith pointed out that the genetic code divides sense and antisense amino acids into three separate groups. In each group, the amino acids and possible antisense amino acids form a network defined by possible base pairings and exchanges due to the redundancy of the genetic code. The significance of this is that, with these groups, amino acids are also classified according to their tendency to form helices, beta sheets or beta turns. This suggests that the DNA complementary strand codes for peptides with similar secondary structural characteristics to the corresponding peptides. In our context, this would mean that similar secondary structures may be assigned to the corresponding antisense boxes, i.e., that the complementary antisense boxes are homologous in their secondary structures. If antisense homology boxes play a significant role in folding and maintaining the structural integrity of proteins, synthesis of such protein fragments could provide peptides that would influence the biological activity of those proteins. We have already found a significant set of experimental examples of interacting antisense homology box derived peptides in C5a anaphylatoxin, C5a receptor and endothelin receptor (a total of 9 out of 13 peptides tested to date exert agonist or antagonist effects on C5a- or endothelin receptor functions [1].

References

[1] Baranyi, L, Campbell, W. & Okada H (1995) Antisense Homology box: a new motif in proteins that encode biologically active peptides. Nature Medicine 9,814-901.

[2] Blalock, J.E. & Smith, E.M. Hydropathic anti-complementarity of amino acids based on the genetic code. BBRC 121, 203-207 (1984).

[3] Blalock, J.E. Complementarity of peptides specified by 'sense' and 'antisense' strands of DNA. TIBTECH 6, 140-144 (1991).