Vol. 121, No. 1, 1984 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 31, 1984 Pages 203-207

Hydropathic Anti-Complementarity of Amino Acids Based on the Genetic Code

J. Edwin Blalock and Eric M. Smith

Department of Microbiology
University of Texas Medical Branch
Galveston, Texas 77550

Received April 17, 1984

An interesting pattern in the genetic code has been discovered. Codons for hydrophilic and hydrophobic amino acids on one strand of DNA are complemented by codons for hydrophobic and hydrophilic amino acids on the other DNA strand, respectively. The average tendency of codons for "uncharged" (slightly hydrophilic) amino acids is to be complemented by codons for "uncharged" amino acids.

While considering the consequences of the transcription of complementary strands of DNA in the 5' to 3' direction and in the same reading frame an interesting pattern in the genetic code was observed. Codons for hydrophilic and hydrophobic amino acids on one strand are generally complemented by codons for hydrophobic and hydrophilic amino acids on the other strand, respectively. The average tendency of codons for "uncharged" (slightly hydrophilic) amino acids was to be complemented by codons for "uncharged" amino acids. This article will detail this pattern.

Complementary DNA Strand Amino Acids

The anti-complementarity of amino acids based on the genetic code is most obvious when presented in terms of the hydropathic character of amino acids. By employing the hydropathic scoring system of Kyte and Doolittle (1), amino acids were assigned a numerical value based on their relative hydrophobic (+ side) or hydrophilic (- side) nature. The hydropathic scale extends from +4.5 (isoleucine) to -4.5 (arginine) and allows for the grouping of amino acids as hydrophobic (+4.5 to +1.8, Table 1), hydrophilic (-4.5 to -3.2, Table 2) and slightly hydrophilic (-0.4 to -1.6, Table 3). Table 1 presents the

0006-291X/84 $1.50

Table 1. Hydropathic Scores for the Amino Acids Whose Codons are Complementary to Those of the Hydrophobic Amino Acids

Coding Strand			Noncoding Strand				
Codon	Amino Acid	Hydropathic Score	Codon	Amino Acid	Hydropathic Score		Average Change
					Individual	Average	
AUU	Isoleucine	+4.5	AAU	Asparagine	-3.5	-2.8	-7.3
AUC	Isoleucine		GAU	Aspartic acid	-3.5		
AUA	Isoleucine		UAU	Tyrosine	-1.3		
GUU	Valine	+4.2	AAC	Asparagine	-3.5	-2.9	-7.1
GUC	Valine		GAC	Aspartic acid	-3.5		
GUG	Valine		CAC	Histidine	-3.2		
GUA	Valine		UAC	Tyrosine	-1.3		
CUU	Leucine	+3.7	AAG	Lysine	-3.9	-3.6	-7.3
CUC	Leucine		GAG	Glutamic acid	-3.5		
UUG	Leucine		CAA	Glutamine	-3.5		
CUG	Leucine		CAG	Glutamine	-3.5		
UUU	Phenylalanine	+2.7	AAA	Lysine	-3.9	-3.7	-6.4
UUC	Phenylalanine		GAA	Glutamic acid	-3.5		
UGU	Cysteine	+2.5	ACA	Threonine	-0.7	+0.6	-1.9
UGC	Cysteine		GCA	Alanine	+1.8		
AUG	Methionine	+1.9	CAU	Histidine	-3.2	-3.2	-5.1
GCG	Alanine	+1.8	CGC	Arginine	-4.5	-0.8	-2.6
GCU	Alanine		AGC	Serine	-0.9		
GCC	Alanine		GGC	Glycine	-0.4		
GCA	Alanine		UGC	Cysteine	+2.5		

codons for the hydrophobic amino acids and their hydropathic scores followed by the complementary codons (noncoding strand) and their corresponding amino acids. Of the possible 20 complementary codons, there are only 2 which code for hydrophobic amino acids. These

Table 2. Hydropathic Scores for the Amino Acids Whose Codons Are Complementary to Those of the Hydrophilic Amino Acids

Coding Strand			Noncoding Strand				
Codon	Amino Acid	Hydropathic Score	Codon	Amino Acid	Hydropathic Score		Average Change
					Individual	Average	
CGC	Arginine	-4.5	GCG	Alanine	+1.8	-0.5	+4
CGU	Arginine		ACG	Threonine	-0.7		
CGA	Arginine		UCG	Serine	-0.9		
AGA	Arginine		UCU	Serine	-0.9		
CGG	Arginine		CCG	Proline	-1.6		
AGG	Arginine		CCU	Proline	-1.6		
AAG	Lysine	-3.9	CUU	Leucine	+3.7	+3.2	+7.1
AAA	Lysine		UUU	Phenylalanine	+2.7		
AAU	Asparagine	-3.5	AUU	Isoleucine	+4.5	+4.4	+7.9
AAC	Asparagine		GUU	Valine	+4.2		
GAU	Aspartic acid	-3.5	AUC	Isoleucine	+4.5	+4.4	+7.9
GAC	Aspartic acid		GUC	Valine	+4.2		
CAA	Glutamine	-3.5	UUG	Leucine	+3.7	+3.7	+7.2
CAG	Glutamine		CUG	Leucine	+3.7		
GAG	Glutamic acid	-3.5	CUC	Leucine	+3.7	+3.2	+6.4
GAA	Glutamic acid		UUC	Phenylalanine	+2.7		
CAC	Histidine	-3.2	GUG	Valine	+4.2	+3.1	+6.3
CAU	Histidine		AUG	Methionine	+1.9		

correspond to the UGC for cysteine and the GCA for alanine on the coding strand. Of the remaining 18 possibilities, 13 were hydrophilic and 5 were slightly hydrophilic. The direction and average change in hydropathic scores of the amino acids of the complementary codons for isoleucine, valine, leucine, phenylalanine, cysteine, methionine and alanine were -7.3, -7.1, -7.3, -6.4, -1.9, -5.1 and -2.6, respectively.

An opposite pattern was observed for the hydrophilic amino acids (Table 2). Of the possible 18 complementary codons on the noncoding strand, none code for strongly hydrophilic amino acids. Thirteen of the 18 were for hydrophobic amino acids, while five of the 18 codons were for slightly hydrophilic amino acids and all of these were complementary codons for a single amino acid (arginine). The direction and average change in hydropathic scores of the amino acids of the complementary codons for arginine, lysine, asparagine, aspartic acid, glutamine, glutamic acid and histidine were +4, +7.1, +7.9, +7.9, +7.2, +6.4, and +6.3, respectively. The pattern for the slightly hydrophilic amino acids differed from that of the two previous groups (Table 3). Of the possible 20 complementary codons, 10 code for slightly hydrophilic and 5 for strongly hydrophilic amino acids. The 5 complementary codons for strongly hydrophilic amino acids are balanced by 5 codons for hydrophobic amino acids. Thus the net effect is for little change in the average hydropathic character of the noncoding strand amino acids. The direction and average change in hydropathic scores of glycine, threonine, tryptophan, serine, tyrosine and proline were 0, -0.1, -0.5, -0.5, +5.8 and -0.9. Tyrosine was the single amino acid in this group which showed a major average change in hydropathic scores of its noncoding strand amino acids. Figure 1 shows a plot of the hydropathic scores of the coding strand amino acids versus the average hydropathic scores of the noncoding strand amino acids. A linear regression analysis of this

Table 3. Hydropathic Scores for the Amino Acids Whose Codons Are Complementary to Those of the Slightly Hydrophilic Amino Acids

Coding Strand			Noncoding Strand				
Codon	Amino Acid	Hydropathic Score	Codon	Amino Acid	Hydropathic Score Individual	Hydropathic Score Average	Average Change
GGU	Glycine	-0.4	ACC	Threonine	-0.7		
GGA	Glycine		UCC	Serine	-0.9		
GGG	Glycine		CCC	Proline	-1.6	-0.4	0
GGC	Glycine		GCC	Alanine	+1.8		
ACC	Threonine	-0.7	GGU	Glycine	-0.4		
ACU	Threonine		AGU	Serine	-0.9		
ACG	Threonine		CGU	Arginine	-3.5	-0.6	-0.1
ACA	Threonine		UGU	Cysteine	+2.5		
UGG	Tryptophan	-0.9	CCA	Proline	-1.6	-1.6	-0.5
UCC	Serine	-0.9	GGA	Glycine	-0.4		
AGU	Serine		ACU	Threonine	-0.7		
UCG	Serine		CGA	Arginine	-4.5	-1.6	-0.5
UCU	Serine		AGA	Arginine	-4.5		
AGC	Serine		GCU	Alanine	+1.8		
UAU	Tyrosine	-1.3	AUA	Isoleucine	+4.5	+4.5	+5.8
UAC	Tyrosine		GUA	Valine	+4.2		
CCC	Proline	-1.6	GGG	Glycine	-0.4		
CCA	Proline		UGG	Tryptophan	-0.9	-2.5	-0.9
CCU	Proline		AGG	Arginine	-4.5		
CCG	Proline		CGG	Arginine	-4.5		

data results in a correlation coefficient of -0.77. A similar pattern is observed when calculated by another hydropathic scoring system which has somewhat different values for tryptophan, tyrosine, glutamine and asparagine (data not shown) (2). Thus the noncoding strand hydropathic scores are inversely related to the coding strand and this relationship is not random.

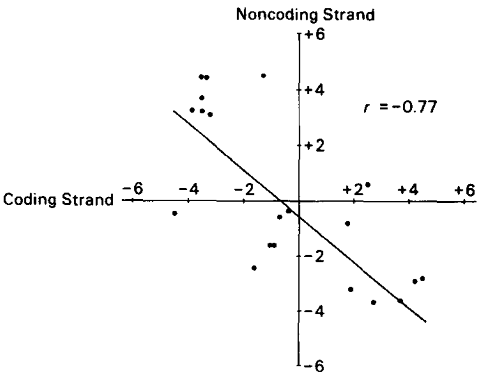

Figure 1. The relationship of the hydropathic scores of the coding strand amino acids and the average hydropathic scores of the corresponding noncoding strand amino acids.

Discussion

An interesting pattern in the genetic code has been observed which allows for hydropathic anti-complementarity of amino acids. Codons for hydrophilic and hydrophobic amino acids on one strand of nucleic acid are generally complemented by codons for hydrophobic and hydrophilic amino acids on the other strand, respectively. The average tendency of codons for "uncharged" (slightly hydrophilic) amino acids was to be complemented by codons for "uncharged" amino acids. This relationship is so striking as to be hardly coincidental and may have an important evolutionary or biological meaning. While we do not presently understand the meaning of the pattern, it is nonetheless intriguing that the complementarity which is a hallmark of nucleic acids results in anti-complementarity in a chemical sense for the amino acids for which they code.

Acknowledgements

We appreciate the encouragement of our colleagues in the Department of Microbiology and particularly Dr. Gerald M. Fuller and Dr. G. John Stanton. This research was supported by a grant from the state of Texas.

References

1. Kyte, J. and Doolittle, R.F., J. Mol. Biol., 157, 105-132 (1982).
2. Hopp, T.P. and Woods, K.R., Proc. Natl. Acad. Sci. U.S.A., 78, 3824-3828 (1981).