RNA (2001), *7*:942–957. Cambridge University Press. Printed in the USA.

Analysis of codon:anticodon interactions within the ribosome provides new insights into codon reading and the genetic code structure

VALERY I. LIM[1] and JAMES F. CURRAN[2]

[1]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 117984 Moscow, Russia
[2]Department of Biology, Wake Forest University, Winston-Salem, North Carolina 27109, USA

ABSTRACT

Although the decoding rules have been largely elucidated, the physical-chemical reasons for the "correctness" of codon:anticodon duplexes have never been clear. In this work, on the basis of the available data, we propose that the correct codon:anticodon duplexes are those whose formation and interaction with the ribosomal decoding center are not accompanied by uncompensated losses of hydrogen and ionic bonds. Other factors such as proofreading, base–base stacking and aminoacyl–tRNA concentration contribute to the efficiency and accuracy of aminoacyl–tRNA selection, and certainly these factors are important; but we suggest that analyses of hydrogen and ionic bonding alone provides a robust first-order approximation of decoding accuracy. Thus our model can simplify predictions about decoding accuracy and error. The model can be refined with data, but is already powerful enough to explain all of the available data on decoding accuracy. Here we predict which duplexes should be considered correct, which duplexes are responsible for virtually all misreading, and we suggest an evolutionary scheme that gave rise to the mixed boxes of the genetic code.

Keywords: codon reading and misreading rules; frameshifting; genetic code; proofreading; RNA structure; translation

INTRODUCTION

Although we know which anticodon:codon complexes are recognized as "correct," we have never understood why only they are acceptable. Crick (1966), based on the emerging structure of the genetic code and base-pair stereochemistry, proposed his famous wobble rules for identifying correct duplexes. He proposed that only canonical base pairing should occur at the first and second codon positions, and that certain wobble pairing would be possible at the third codon position. In succeeding years these general rules have been amply confirmed, although the range of acceptable wobble pairs has been expanded (Osawa et al., 1992; Boren et al., 1993; Inagaki et al., 1995). There has also been progress towards an understanding of how nucleoside modifications affect wobbling (e.g., Agris, 1991; Björk, 1992, 1998; Osawa et al., 1992; Yokoyama & Nishimura, 1995; Curran, 1998). However, the physical-chemical properties that underlie these rules for correct codon:anticodon duplexes have never been clear.

The ultimate determinant of aminoacyl–tRNA selection must be codon:anticodon stability. But the stabilities predicted from solution studies of nucleic acid interactions do not reliably distinguish the correct duplexes from the incorrect ones. Stable RNA double helices can contain a wide variety of mismatches and even blocks of mismatches (e.g., Holbrook et al., 1991; Baeyens et al., 1995; Dirheimer et al., 1995). Mismatches and their blocks are easily incorporated into helices by virtue of the mobility of the polynucleotide chain in the vicinity of the A-form conformation. Stable but wrong anticodon:anticodon duplexes with the UU pair in the middle of the minihelix are also observed in the crystals of yeast $tRNA^{Asp}$ (Moras et al., 1980). Moreover, the minihelices formed in buffer by two tRNA anticodons, which are the best available models for codon:anticodon stability because those duplexes neatly control for anticodon loop features such as nucleoside

Reprint requests to: James F. Curran, Department of Biology, Wake Forest University, Winston-Salem, North Carolina 27109, USA; e-mail: curran@wfu.edu.

Definitions: a_1, a_2, a_3 and c_1, c_2, c_3: the anticodon (*a*) and codon (*c*) nucleotides; a_1c_3, a_2c_2, a_3c_1: the codon:anticodon duplex pairs, for example, U_3G_1 is the pair formed by G in the first codon position and by U in the third anticodon position; P_Y and P_U: pyrimidine and purine bases. SREs: Steric restriction elements that cause uncompensated losses of hydrogen and ionic bonds at the formation of codon:anticodon duplexes and their interaction with the ribosome.

modification and stereochemical constraints, contain mismatches that are not allowed in correct duplexes (Grosjean et al., 1978). In fact, some of these mispaired complexes are just as stable as duplexes that contain only correct base pairs. Clearly, both correct and wrong codon:anticodon duplexes can be stable in solution. Notice that ribosomal proofreading, which can in principle amplify small energetic differences (Hopfield, 1974; Ninio, 1975; Kurland et al, 1990; Yarus, 1992), cannot distinguish duplexes that have essentially the same stabilities. Therefore, in addition to using a proofreading mechanism, ribosomes must rely on features other than duplex stability as predicted from solution and structural studies.

There is direct evidence that the ribosomal decoding center strongly distinguishes complexes that should have similar solution stabilities. Even when proofreading is inhibited, ribosomes programmed with the UUU triplet can distinguish the correct $tRNA^{Phe}$ from the near cognate $tRNA_2^{Leu}$ by a factor of $10^{-3.5}$ (Thompson & Karim, 1982), which implies an energetic difference of at least 20 kJ/mol. The principal difference between them is that the cognate duplex has a first-position AU pair and the near-cognate has GU. Even with a generous allowance for stronger stacking for the cognate due to nucleoside modifications in the anticodon loop, one would predict a much smaller difference between these duplexes.

One of us (Lim & Venclovas, 1992) and others (Potapov, 1995; Burkhardt et al., 1998) have previously suggested that ribosomes place stereochemical constraints on duplex structure such that only correct duplex-decoding center complexes are stable. Here, we extend this idea to develop a model for duplex recognition that is consistent with all data on decoding accuracy and with the emerging structural models for the ribosomal decoding center. We consider the types of interactions and stereochemical conditions of the ribosomal decoding center that could provide for correct codon reading. Stereochemical and thermodynamic analyses demonstrate that the non-deformed A-form of the codon moiety of the duplexes should be the structural invariant whereby the decoding center recognizes correct duplexes. This model allows, for the first time, an explanation of why the correct duplexes work, including an explanation of the wobble rules. As a corollary the model also provides rules for misreading that are consistent with the available data on misreading. Furthermore, a particularly robust feature of the model predicts that the duplex assembly and recognition occur separately and in that order, which is consistent with observations that ribosomal conformational changes occur during tRNA-ribosomal association. We also discuss how these results relate to the roles of proofreading and tRNA nucleoside modification on accuracy. Our results also illuminate the rationale for the structure of the split boxes and their distribution in the genetic code dictionary. They support the startling notion that the split boxes arose from family boxes that were functionally divided to prevent the corresponding tRNAs from misreading at the *second* codon position.

THEORETICAL BACKGROUND

Three propositions form the basis of our stereochemical analysis.

The first proposition

The nondeformed A-form of the codon moiety is the structural invariant in the correct duplexes. The ribosomal decoding center distinguishes the correct from wrong duplexes by the recognition of this invariant.

The second proposition

A loss of a single hydrogen or ionic bond at the duplex formation or its interaction with the decoding center is sufficient to distinguish correct from wrong duplexes if the lost bond cannot be compensated by the formation of alternate bonds. These uncompensated losses of the bonds are caused by steric restriction elements (SREs: separate atoms, bases, amino acid side chains, etc.). Lost bonds provide the required difference of about 20 kJ/mol or more in the free energy to distinguish correct and wrong codon reading.

The third proposition

All of the steps of the ribosomal cycle must be more rapid than the cycle itself. The disruption of hydrogen and ionic bonds imposes energetic barriers that must be surmounted during translation. We show that the number, N, of hydrogen and/or ionic bonds that must be simultaneously broken plus the number, M, of uncompensated lost bonds must be less than four (i.e., $N + M < 4$) to avoid kinetic barriers that would be insurmountable within the typical ribosomal cycle ($\sim 10^{-1.9}$ s; Nierhaus, 1993). This fact imposes strong constraints on the mechanism of duplex assembly and disruption as well as the interaction and dissociation of the duplex with the decoding center. This fact must also affect other steps such as translocation and so forth, but here we restrict our analysis to the duplex recognition steps.

Let us consider experimental data that lead to the above propositions and general stereochemical restrictions following from these propositions.

Proposition 1

The decoding center must recognize some structural invariant of all correct duplexes regardless of whether codons are paired with tRNA anticodons or release

factor "anticodons." Because of large stereochemical differences between polynucleotide and polypeptide chains, only the codon moiety can contain a structural invariant specifically recognized by the decoding center. This is demonstrated well by the recently determined (Song et al., 2000) crystal structure of release factor. Although domain 1 has a gross structural similarity to the tRNA anticodon arm, microscopically the protein anticodon looplike section is an α-helical hairpin that differs drastically from the crystal tRNA anticodon loop structure in the distribution of polar and nonpolar atoms, in size and form. Therefore, the decoding center must direct its specific recognition to the codon moiety of the duplex. Moreover, the decoding center should specifically recognize the nondeformed A-form conformation of the codon, which is the only single codon conformation compatible with all cognate duplexes.

As for the anticodon, the structure of the tRNA anticodon loop imposes its own limitations on the anticodon conformation. The mobility of the first anticodon base a_1 in tRNAs is high whereas the mobility of a_2 and a_3 is sterically restricted in the vicinity of the A-form (Lim & Venclovas, 1992). The required codon A-form and the low mobility of a_2 and a_3 counteract duplexes with noncanonical base pairing in the pairs a_3c_1 and a_2c_2. At the same time the mobility of the anticodon a_1 allows restricted wobble in the pair a_1c_3.

Proposition 2

As noted in the Introduction, even when proofreading is inhibited, ribosomes can discriminate between cognate and near cognate complexes by factors of $\sim 10^{-3.5}$, which implies a free energy difference of $\geq$20 kJ/mol (20 kJ/mol at equilibrium, which must be the minimum difference; for background on thermodynamic calculations, see Materials and Methods). Codon:anticodon duplexes are stabilized by hydrogen bonds, ionic interactions, and base stacking. Which of these can be responsible for distinguishing correct from wrong duplexes? Experiment shows that base stacking is wholly inadequate. Even the complete disruption of stacking between two bases (removal of a base from a helix) gives the free energy change that does not exceed 4 kJ/mol (Ts'o, 1974), which is far less than the required $\sim$20 kJ/mol. In the average these changes are on the order of the thermal energy kT (2.5 kJ/mol). Furthermore, that a wide variety of mismatches is tolerated in aqueous solution, in which the effects of base stacking are maximized, also makes it clear that changes in base stacking are not able to distinguish wrong and correct duplexes.

Note that, although stacking interactions cannot be used as the critical determinants of duplex correctness, changes of stacking interactions could affect translation. They may contribute to distinctions among the various *correct* duplexes, for example. The efficiencies of correct duplexes formed by the anticodon with synonymous codons may vary by several-fold relative to each other when the differences in the base stacking energy vary by a few kiloJoules per mole. However, it is abundantly clear that differences in base stacking cannot provide for the differences in efficiency of several orders of magnitude that separate correct from errant reading.

Alternatively, the disruption of a hydrogen or ionic bond, if it is not replaced by another such bond (i.e., an uncompensated loss of the bond), can provide the necessary free energy change of $\sim$20 kJ/mol between correct and wrong complexes (20 $\pm$ 5 kJ/mol for disruption of a normal hydrogen bond, slightly more for an ionic bond; Pauling & Pauling, 1975; Saenger, 1984). The requirement that the broken hydrogen or ionic bond is uncompensated is important because if the disrupted bond is compensated by new bonds, then the full $\sim$20 kJ/mol increase in free energy will not be realized. To illustrate this point, consider the formation of an RNA duplex. Duplex formation is always accompanied by the substitution of one set of hydrogen and ionic bonds for another. Prior to duplex formation all polynucleotide polar atoms form such bonds, either internally or to solvent molecules. During formation of the duplex, some of these bonds are disrupted and replaced by others, including (but not limited to) base–base bonds. As discussed in the Introduction, mismatches and so forth do not always destabilize RNA duplexes in solution. That is because in those cases, the posthybridization bonds compensate for the prehybridization bonds lost during duplex formation. To distinguish cognate from errant duplexes the ribosome must, therefore, sterically restrict duplexes such that only cognate complexes form fully compensating bonds in the decoding center. This would lead to an increase in the enthalpy part of its free energy by about 20 kJ/mol per uncompensated, disrupted bond in mispaired duplexes. Because hydrogen and ionic bonds have the requisite energy, the uncompensated loss of a single such bond is adequate to cause rejection of incorrect duplexes.

SREs are required to provide uncompensated losses of hydrogen and ionic bonds. SREs should prevent the formation of new hydrogen and ionic bonds in exchange for the disrupted ones. The role of such an SRE can be played by separate atoms (Fig. 1), bases (Figs. 2 and 3), amino acid side chains, and so forth. Below we argue that the ribosome must also use SREs to provide uncompensated losses of hydrogen and ionic bonds at the interaction of wrong duplexes with the ribosomal decoding center.

Proposition 3

Because high energies are required to disrupt them, hydrogen and ionic bonds can also create significant

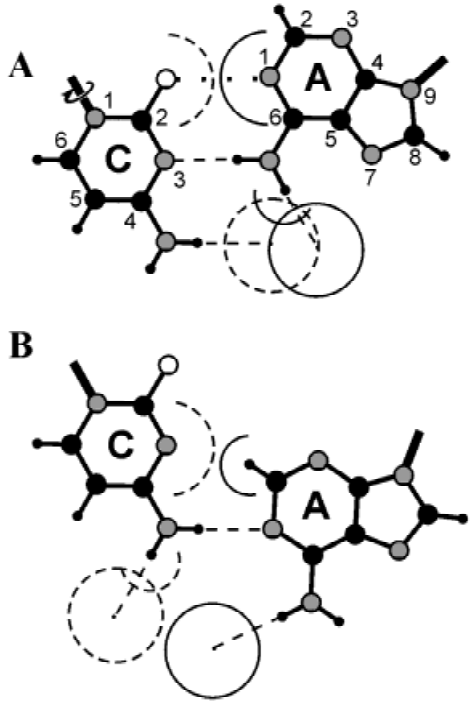

FIGURE 1. AC pairing with a coplanar arrangement of the bases. **A:** GU-like pairing, the mutual orientation of bases is identical to that in GU. Extrabold lines are glycosyl bonds. Arcs and circles demonstrate the van der Waals sizes of atoms. N3 of C and $(NH_2)6$ of A form a base–base hydrogen bond. Dotted line is a base–base hydrogen bond that could be formed by protonation of N1 of A or by replacement of O2 of C by a hydrogen bond donor. The atom O2 of C plays the role of an SRE that prohibits the bond of N1 of A with solvent. The other SRE is $(NH_2)6$ of A. The oxygen of a water molecule (broken circle) cannot be hydrogen-bonded to $(NH_2)4$ of C without a very strong steric overlap with $(NH_2)6$ of A and its water molecule (the second circle). The action of SREs can be eliminated by the nonstandard propeller twist formed by a positive rotation of the base (an arrow) around the glycosyl bond (for details, see Materials and Methods). **B:** Alternative AC pairing with the only base–base hydrogen bond $(NH_2)4(C)\bullet\bullet\bullet N1(A)$. The only SRE (H2 of A) shields N3 of C from solvent.

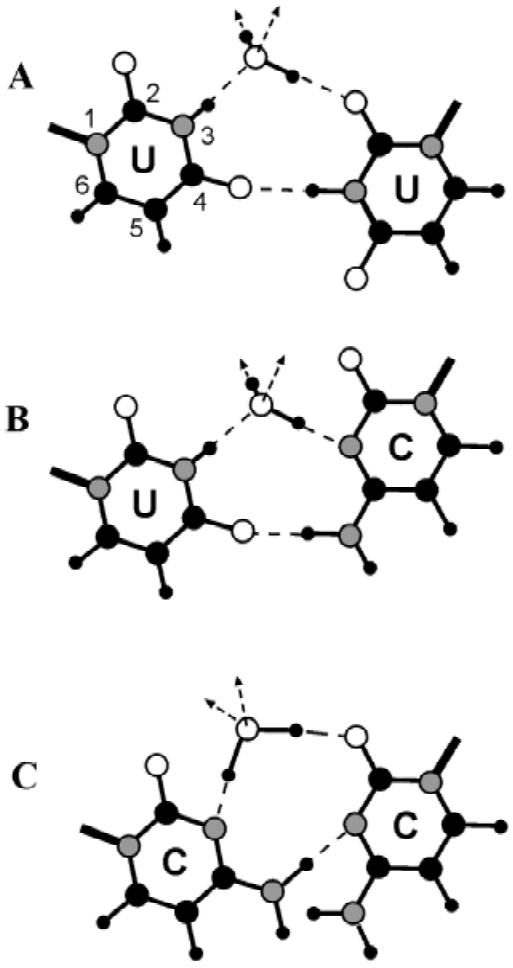

FIGURE 2. Pyrimidine:pyrimidine pairs $P_Y\text{^}P_Y$ containing a base–water–base bridge. **A:** The pair U^U. **B:** The pair U^C. The bases in the pairs are coplanar. Extrabold lines are glycosyl bonds. A water molecule can form four hydrogen and ionic bonds that are approximately tetrahedrally oriented. A bridging water molecule in the pairs U^U and U^C forms two bonds with the bases. The other two bonds (arrows) are oppositely directed from the base-pair plane. For this reason, their formation in the pairs located in c_1 and c_2 is counteracted by adjacent base pairs (see Fig. 3). **C:** This figure demonstrates that the pair C^C cannot be formed for steric reasons. Simultaneous formation of a water bridge and base–base hydrogen bond in this pair is prohibited by a very strong steric overlap between the cytosine NH_2 groups.

kinetic barriers in translation. The height of a barrier to disruption of a hydrogen or ionic bond approaches the energy of the breaking bond (~20 kJ/mol), because for steric reasons the formation of a new bond(s) can occur only after practically full disruption of the original bond. When more than one hydrogen and/or ionic bonds must be simultaneously disrupted, the barrier is 20 kJ/mol times the sum of the bonds; that is, barriers can become very high as the number of bonds increases. These barriers confront the formation and disruption of duplexes, as well as the dissociation and association of duplexes with the ribosome. For example, the formation of base–base hydrogen bonds is a replacement of base–water hydrogen bonds, and this displacement is confronted by the barrier to the disruption of the base–water bonds. Duplex disruption is confronted by the reverse reaction.

In the general case, the time required to overcome an energetic barrier(s) caused by hydrogen and ionic bonds is determined by the barrier height of $(N + M) \times 20$ kJ/mol (Fig. 4), where N is the number of simultaneously disrupted bonds, and M is the number of uncompensated, lost bonds in the complex. Clearly, the number $N + M$ cannot impose a kinetic barrier that would preclude the formation and disruption of duplexes or the interaction and dissociation of duplexes with the decoding center within the normal time of the ribosomal cycle ($\sim 10^{-1.9}$ s; e.g., Nierhaus, 1993). The average time required to overcome barriers imposed by $N + M = 3$ (3×20 kJ/mol) is $\sim 10^{-2.5}$ s (for background on the calculations, see Materials and Methods), which is much faster than the typical ribosomal cycle. In contrast, the average time required to over-

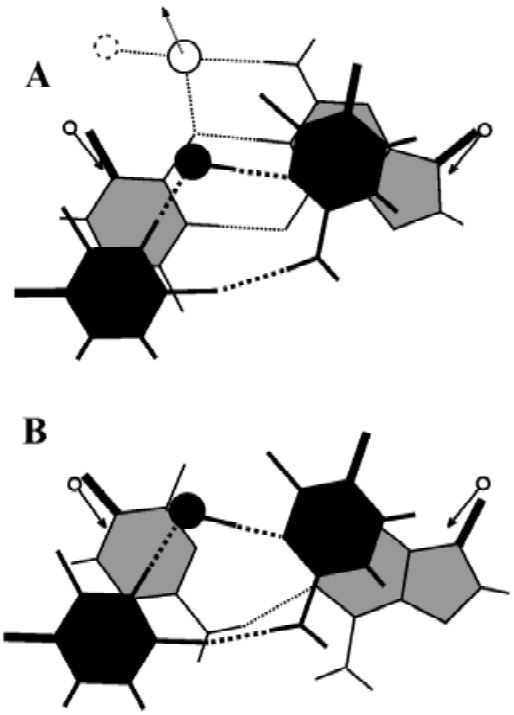

FIGURE 3. The wobble pair U_1^C_3 and the pairs C_2A_2 and U_2G_2 with quasi-canonical orientations of the glycosyl bonds. **A:** The pair U_1^C_3 (bold, presented towards the reader) and the pair U_2G_2 (fine, positioned under U_1^C_3). **B:** The pair U_1^C_3 and the pair C_2A_2. At the left and right of **A** and **B**, respectively, the anticodon and codon bases are shown. Bases in all four pairs are coplanar. To demonstrate a quasi-canonical orientation of the glycosyl bonds in C_2A_2 and U_2G_2, positions of the glycosyl bonds in the canonical pairs (arrows with small circles in the immediate vicinity of the glycosyl bonds of U_2G_2 and C_2A_2) are given. An open circle in **A** is a water molecule simultaneously hydrogen-bonded to NH_2 of G, the base atom O2 of U (gray circle) and its ribose (OH)2′ group (broken circle). The fourth possible bond of this water molecule is directed towards the edge of the wobble pair a_1c_3. When a_1c_3 is not P_Y^P_Y, this fourth bond with solvent cannot be realized because of a steric restriction created by the edge of a_1c_3, that is, the edge of the wobble pair plays the role of SRE. The other pair (a_2c_2) also plays the role of SRE. It prohibits the bond of bridging water molecules (black beads) directed towards its edge (for this bond, see Fig. 2).

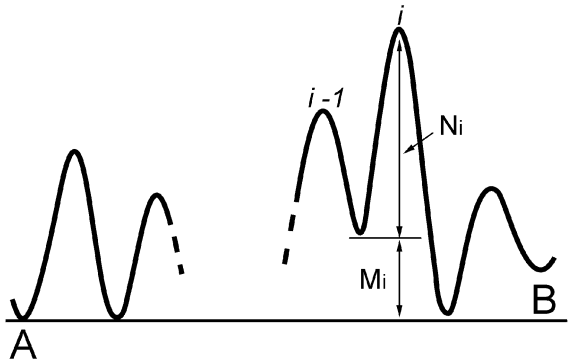

FIGURE 4. Energetic barriers caused by disruption and formation of hydrogen and ionic bonds during transition of some structure from state A to state B. The barrier heights are expressed in terms of the energy of a hydrogen bond. The height of the ith intermediate barrier relative to state A is $N_i + M_i$. Here, N_i is the number of simultaneously disrupted bonds forming the left slope of the ith intermediate barrier, and M_i is the difference in the number of uncompensated losses of the bonds between the initial state A and after overcoming the $(i-1)$th barrier. The transition time of a structure from state A to state B is determined by the maximum intermediate barrier relative to state A, that is, by the maximum value of $N_i + M_i$.

come a barrier of 4×20 kJ/mol is ~10 s. Because this time exceeds that of the ribosomal cycle by 2.9 orders of magnitude, no step of the translational cycle can face such a barrier. Below, we discuss these facts as rule $N + M < 4$.

RESULTS

Stereochemical and kinetic considerations show that the decoding center cannot simply bind all mobile atoms of the codon to lock it in the A-form

Although one might imagine that the ribosome could simply form many hydrogen and ionic bonds to fix the A-form of the codon, stereochemical considerations show that this is not possible. According to the rule $N + M < 4$, no more than three hydrogen and ionic bonds recognizing the codon A-form can be simultaneously formed (or disrupted) between the decoding center and the duplex. Therefore, the duplex cannot simply move next to a ribosomal surface or "pocket" that simultaneously establishes bonds to all mobile atoms. Formally, it would be possible to increase the number of bonds to four or more if some are formed and disrupted sequentially. The sequential breaking of two bonds, for example, could occur by the rotation of the duplex or the decoding center around one of the bonds. Such rotation can break the radial bond while leaving intact the axial bond, which could then be broken in turn. This sequential breaking of bonds would allow each step to be confronted only by the kinetic barrier due to the breakage of the respective single bond.

But when two bonds are in close proximity, a large rotation is required to disrupt the radial bond. The small distances between bonds within the codon would require rotations that are accompanied either by large shifts (several tens of Angstroms) of different parts of tRNA relative to the ribosome, or by significant rearrangements of the decoding center inside the ribosome. Such large shifts of tRNA are not compatible with the crystal structures of 70S ribosome (Cate et al., 1999) and the high-resolution 30S crystal structure (Carter et al., 2000). As for rearrangements of the decoding center, they cannot be done without breakage of the structural domains in the neck region of the 30S subunit. Because such rotations are disallowed, to avoid violating the rule $N + M < 4$ the decoding center should use no more than three bonds to recognize the codon A-form.

However, three duplex-decoding center bonds are not adequate to fix the codon in any particular conformation. Therefore, the codon A-form must also be stabilized by intraduplex bonds. Below we argue that the participant bonding groups do not have access to alternative pairing partners and that the disruption of at least one such bond occurs for every wrong duplex. This system neatly provides for the distinction of the correct and incorrect duplexes based on fundamental features of the RNA A-form: the interribose hydrogen and ionic bonds.

Codon interribose hydrogen and ionic bonds together with decoding center SREs acting on these bonds should provide for recognition of the correct duplexes

The ribose (OH)2′ group possesses both hydrogen bond donor and acceptor properties (Gurskaya, 1968; Jeffrey et al., 1985). Therefore it simultaneously forms two coplanar hydrogen and ionic bonds that are located approximately in the plane C2′O2′H2′. Having both donor and acceptor properties, (OH)2′ allows the organization of two types of bonding systems that could fix the A-form conformation of the codon in a sequence-independent manner. One system is interribose hydrogen bonding (formally "hydrogen" bridging) between the O2′ and O4′ of adjacent ribose rings (Fig. 5). The other system is interribose cation bridging between the same atoms. A bridging cation forms two additional bonds in a tetrahedral arrangement (Fig. 5). Either type of bonding system organizes the RNA backbone into five linked rings, and these rigidly linked rings fix the codon A-form, and deviations from the A-form caused by the wrong wobble pairs and mismatches in c_1 and c_2 disrupt these interribose bonds.

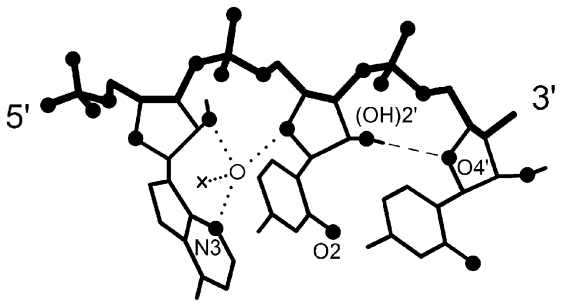

FIGURE 5. An RNA fragment $P_UP_YP_Y$ in the A-form conformation. Black beads are the polar atoms including N3 in P_U and O2 in P_Y (they are practically identically located relative to the glycosyl bond) that can form hydrogen and ionic bonds sequence independently. The other polar atoms of $P_UP_YP_Y$ (they are not highlighted) are variously arranged; therefore the formation of bonds by them is sequence dependent. The broken line is the interribose hydrogen bond (OH)2′•••O4′. Dotted lines are the approximately tetrahedrally oriented four bonds formed by the K^+- or Na^+-like cation (the open circle). Two bonds are formed by a cation with the ribose (OH)2′ and O4′ (the interribose cation bridge), the third bond with the base polar atom N3 or O2, the fourth bond with a solvent molecule (x). The interribose hydrogen bond and bonds formed by a cation are positioned at the minor groove surface of a double helix.

To recognize incorrect duplexes, disrupted interribose bonds should not be compensated by alternative hydrogen and ionic bonds. Such uncompensated losses can occur only if the decoding center SREs act on the codon so that the codon riboses are unable to establish alternate bonds. Figure 6A depicts different hypothetical variants of such SREs created by adenines from the decoding center. Note that the model does not require adenines for this role; other bases, amino acids, and so forth could serve as well. One adenine (gray) shields the c_2-c_3 interribose bond from solvent. The second bond of the ribose (OH)2′ of c_2 is formed to a solvent molecule. After disruption of the shielded interribose bond, the ribose O4′ component will not be able to form a new bond with solvent and (OH)2′ will also not be able to form a new pair of bonds. Consequently, a simple shielding of the interribose bond between two adjacent codon residues provides a recognition of the A-form conformation.

The other adenine (bold) is approximately directed across the sugar–phosphate backbone. The ribose (OH)2′ of c_1 forms two bonds: a hydrogen bond with N1 of A and ionic bond with a cation fixing the A-form of c_1 and c_2 (Fig. 6A). Mismatches in c_1 or c_2 disrupt the c_1-c_2 cation bridge and often they also disrupt a hydrogen bond formed by adenine with (OH)2′ of c_1. However, even when mismatches do not touch the adenine–(OH)2′ hydrogen bond, it will be disrupted by rotation of (OH)2′ around its C2′–O2′ covalent bond. The rotation is required so that, after cation bridge disruption, (OH)2′ of c_1 and O4′ of c_2 can form the bonds with solvent. In other words, (OH)2′ of c_1 in duplexes with mismatches at c_1 or c_2 should form two bonds with solvent with a new spatial orientation. However, N1, H2, and $(NH_2)6$ of the SRE adenine sterically prohibits it. Consequently, the adenine hydrogen-bonded to the (OH)2′ group of the *i*th codon ribose (as is shown, e.g., in Fig. 6A) plays the role of SRE for the c_i-c_{i+1} interribose bond.

SREs also destabilize duplexes with bulges or with only 2 bp. In principle, the lack of a decoding center pocket (above) allows duplexes with bulges and with 2 bp (Fig. 7) to be formed. The stability of duplexes with two pairs can be comparable to duplexes with three pairs, and bulges (both small and large) are found in RNA double helices (e.g., Dirheimer et al., 1995; Conn et al., 1999; Wimberly et al., 1999). But the decoding center SREs can prohibit such duplexes. Bulges disrupt the interribose bonds and therefore SREs should also provide uncompensated losses of the bonds. Moreover, the bold and gray adenines show that SREs can prohibit duplexes with 2 bp. These duplexes are obtained by removal of c_1 or c_3 from the minihelix (Fig. 6B). After removal of c_1, the ribose atoms O4′ of c_2 and N1

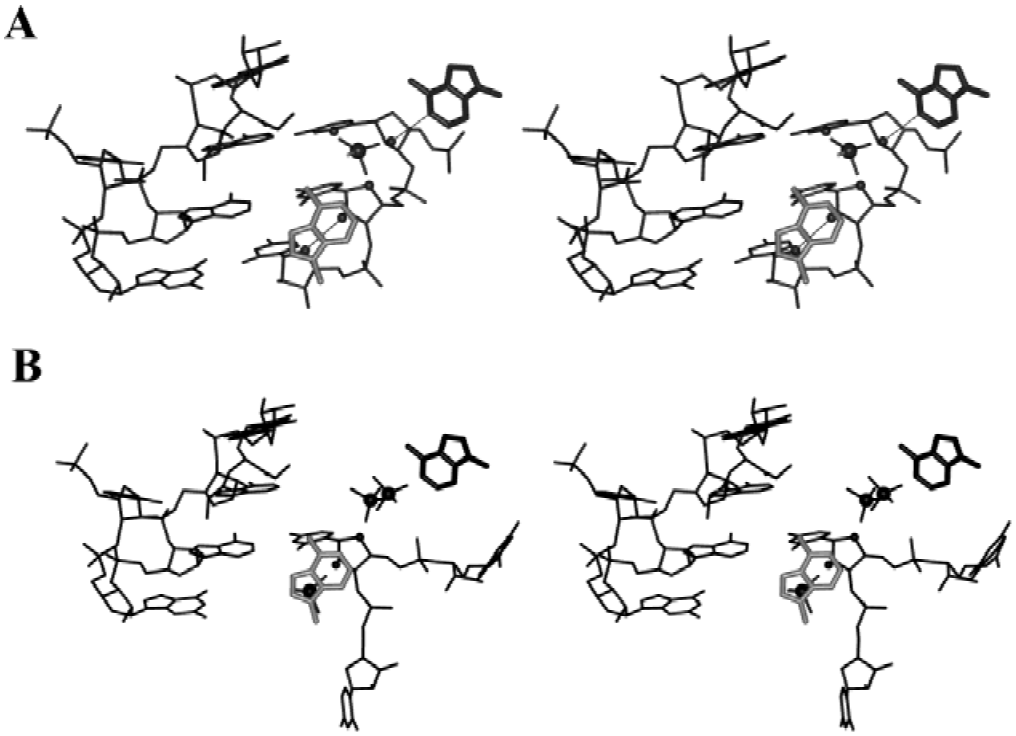

FIGURE 6. SREs formed by adenines from the decoding center that could provide the recognition of the codon A-form. **A:** Stereo view of two different adenine SREs (gray and bold adenines) interacting with the duplex formed by the anticodon loop section 33–37 of $tRNA^{Phe}$ (left) with UUU triplet (right). The orientations of adenines are chosen to avoid bond losses by adenines and duplex polar atoms in the presence of the interribose bonds. Small black beads are O4′ and (OH)2′ of the riboses and O2 of the codon base U_1. A large bead with sticks is a cation interacting with (OH)2′ and O4′ of c_1 and c_2, respectively, and with O2 of the codon base U_1. The hydrogen bond acceptor atom N1 of one adenine (bold) forms a hydrogen bond (fine line) with (OH)2′ of c_1. This hydrogen bond is disrupted at the rotation of (OH)2′ around its covalent bond C2′-O2′. The second adenine (gray) shields the interribose hydrogen bond (fine line) formed by (OH)2′ of c_2 and O4′ of c_3. **B:** Interaction of solvent molecules (three large beads) with O4′ and (OH)2′ (two small beads) of c_2 and with N1 of the bold A in the absence of c_1 or c_3 in the minihelix. Solvent molecule sticks show the tetrahedral arrangement of possible bonds with other solvent molecules.

of the bold adenine should interact with solvent molecules in exchange for the disrupted c_1-c_2 cation bridge and the hydrogen bond N1(A)•••(OH)2′(c_1). However, for steric reasons the O4′ and N1 cannot simultaneously interact with solvent without a loss of at least one bond (Fig. 6B). After removal of c_3, the gray A also sterically reduces the number of bonds that can be formed by a solvent molecule bonded to (OH)2′ of c_2 (Fig. 6B) in the absence of the gray A.

Hence, we can conclude that the recognition of the codon A-form in the codon:anticodon duplexes should be provided by codon interribose bonds together with their SREs. Strong support for the determining role of the codon interribose bonds in the recognition of the correct duplexes has been obtained in the work of Potapov et al. (1995). These authors revealed that the A-site codon lacking the (OH)2′ groups (DNA codon) is not accepted and prevents occupation of the A site.

As to the three allowed decoding center-duplex bonds recognizing the codon A-form, probably these bonds also exist. For example, such a bond could be the hydrogen bond formed by the bold adenine (Fig. 6A).

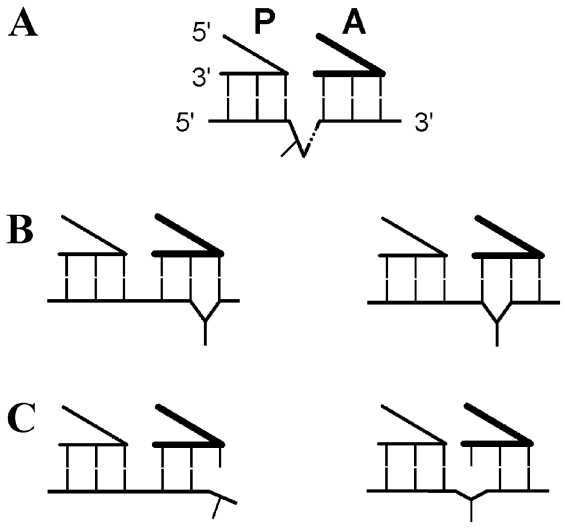

FIGURE 7. Patterns of possible codon:anticodon duplexes. Acute angles are the tRNA anticodon loops. The anticodon loop of the P-site tRNA is shown in fine lines. Bold lines are the anticodon loop of the A-site tRNA. **A:** Bulges in mRNA between the P- and A-site duplexes. **B:** The A-site duplexes with bulges. **C:** The A-site duplexes with two base pairs.

Besides providing an extra contribution into the recognition of the codon A-form, they can fix the codon relative to SREs and can help to prohibit PA interduplex bulges (Fig. 7), which could lead to frameshifting.

Assembly of codon:anticodon duplexes should occur outside of the influence of SREs of the codon interribose bonds to avoid violating the rule $N + M < 4$

The rule $N + M < 4$ imposes strong constraints on duplex formation and disruption in the decoding center. Consider duplex disruption (formation is the reverse reaction; it is simpler to describe disruption). Because anticodons are significantly fixed by anticodon loop structure, and because each base pair contains at least two hydrogen bonds, the avoidance of large kinetic barriers requires that duplex disruption occur via the consecutive removal of the three codon residues from the minihelix. As for base–base hydrogen bonds, low anticodon mobility and flexibility impede the shifts of the glycosyl bonds and base rotation around these bonds needed for sequential breakage. Therefore, AU pairs in the minihelix are generally broken in a single two-bond step. As for GC, the three bonds can be disrupted in steps of one and then two bonds using a nonstandard propeller twist (see Materials and Methods) when the codon interacts with its SREs, or standard twist in the absence of SREs.

Consider the disruption of correct duplexes outside of the influence of the SREs of the codon interribose bonds. In this case, the sequential removal of codon residues from the minihelix will be confronted by barriers created by the simultaneous breakage of only two base–base hydrogen bonds. The codon interribose bonds do not create barriers because they are not required in the absence of SREs. They can be replaced by bonds formed by (OH)2′ and O4′ with solvent. Therefore, in the absence of SREs, disruption of the correct duplexes does not face insurmountable kinetic barriers. This is not true for duplex disruption under the influence of the SREs.

The barrier heights increase in the presence of codon SREs because the SREs make the interribose bonds essential. Thus removal of the central base (c_2) from the minihelix is accompanied by simultaneous breakage of four bonds (two interribose bonds and two base-base bonds; $N = 4$). When the first removed residue is c_1 or c_3, in these cases their removal leads to the duplexes with two base pairs. These duplexes have at least one uncompensated lost bond under the influence of SREs (Fig. 6). Therefore removal of another codon residue after removal of c_1 or c_3 will be accompanied by the formation of a barrier the value $N + M$ of that is equal to $3 + 1 = 4$ (3 is a breakage of two base-base bond plus one interribose bond, 1 is an uncompensated loss of the bond in duplexes with two base pairs). Consequently, we have barriers of $N + M = 4$ for intermediate steps of every conceivable route for the disruption of cognate duplexes in the presence of the SREs.

Similar considerations show that formation of wrong duplexes also cannot occur under the influence of SREs without violating the rule $N + M < 4$. There is one significant difference between the disruptions of correct and wrong duplexes. In contrast to the correct duplexes, the disruption of wrong duplexes should occur faster than their formation because the SREs increases the energetic levels of wrong duplexes by $\sim M \times 20$ kJ/mol.

Thus we see that the formation and disruption of both the correct and wrong duplexes cannot occur under the influence of SREs without violation of the rule $N + M < 4$. This means that duplex assembly should occur outside of the influence of SREs. Following duplex assembly, the SREs and duplexes may be brought into proximity for codon A-form recognition. In addition, up to three decoding center-duplex bonds may form at that time.

Stereochemical model for codon reading

Description of the model

A model (Fig. 8) for codon reading follows naturally from the above considerations. According to this model the selection of the correct duplexes requires at least two stages.

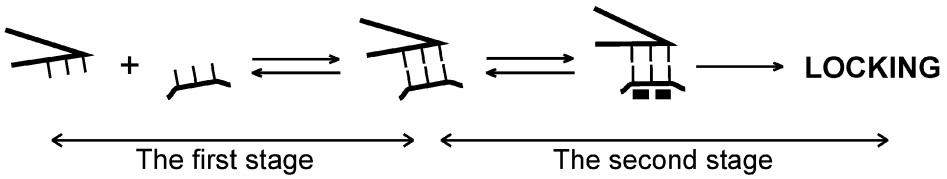

FIGURE 8. A model for codon reading. The acute angle is the tRNA anticodon loop. Its anticodon forms a duplex with the codon (base triplet). Two small, filled rectangles interacting with the codon are SREs of the codon interribose bonds. The first stage is the duplex formation outside of the influence of the SREs. Only the duplexes without uncompensated losses of hydrogen and ionic bonds (stable duplexes) participate in the second stage of the duplex selection. At this stage only stable duplexes that are compatible with SREs are involved into the ribosomal cycle (locking).

The first stage is assembly of the duplex of the aminoacyl–tRNAs outside of the influence of the decoding center SREs of the codon interribose bonds. The various duplexes will have various stabilities. Those that are very unstable due to many mismatches and misalignments will rapidly dissociate. However, because the duplexes are not yet sterically restricted from compensating bonds broken due to mismatches, a variety of stable, especially near-cognate complexes may persist, essentially as predicted from solution studies of RNA duplex stability.

The second stage is the distinction of the correct duplexes from the stable but wrong ones. At this stage the duplex and the decoding center SREs should be drawn together. In principle, the formed duplex could move into the influence of SREs, or SREs could move into position over the preformed duplex, and we cannot distinguish these possibilities. Regardless, as a result of the action of SREs, all correct duplexes will not have uncompensated losses of bonds and will be locked in the decoding center. In contrast, SRE–duplex complexes in which the codon interribose bonds cannot be formed without bond losses in the rest of the duplex will rapidly dissociate.

This model describes, for the first time, a well-formulated structural definition of the correct duplexes, allowing them to be found by a simple stereochemical model. As pointed out above, the mobility of the first anticodon base a_1 is high whereas the mobility of a_2 and a_3 is low in the vicinity of the A-form. At the same time the decoding center should recognize only the duplexes with the nondeformed codon A-form. All of this leads to a simple stereochemical model in which only a_1 is mobile while the other five duplex residues are fixed in the A-form. Only such duplexes without uncompensated losses of hydrogen and ionic bonds should be considered as the correct ones.

The identification of duplexes that contain uncompensated lost hydrogen and ionic bonds can be performed by consideration of only the duplex polar atoms and solvent molecules bonded to them. Such stereochemical analyses can be accomplished using computer graphics techniques (see Materials and Methods).

The wobble rules

Previously, one of us (Lim & Venclovas, 1992; Lim, 1994, 1995) identified the correct wobble pairs a_1c_3 using such a stereochemical modeling. These pairs are presented in Table 1. Besides the base pairs described in the work of Crick (1966), the pairs U^U, U^C, and C^C with a base–base water bridge ("^" symbolizes the water bridge; Fig. 2) have also been used. These pairs have never been considered in the codon:anticodon interactions. We have shown that only such P_YP_Y wobble pairs can be used when the codon is fixed in the A-form (see Materials and Methods).

TABLE 1. Wobble rules for unmodified and modified residues in the first anticodon position.

Nucleoside on the anticodon	Nucleoside recognized on the codon
U	A, G, U, C
C	G
A	A < G < C < U (A and G poorly)
G	U, C
I	U, C, A (A poorly)
S^2U	A, G (G poorly)
Se^2U	A, G
Um	A, G (G less well than A)
xm^5U	A, G
xo^5U	A, G, U
k^2C (lysidine)	A

"Poorly" means that recognition efficiency is several percent. "Less well" means that recognition efficiency is several times lower. A < G < C < U is a rank of recognition efficiency of all the four standard bases.

The wobble rules following from our model (Table 1) are significantly different from those proposed by Crick (1966). In accordance with Crick's rules, the anticodon wobble U recognizes codon A and G (according to our rules, U can also recognize U and C); C recognizes G; A recognizes U (according to our rules, A recognizes all standard bases U, C, A, G); G recognizes U and C; I recognizes U, C, A (according to our rules, inosine should recognize A less well than U and C). Note that some of the wobble pairs may cause backbone distortions and/or poor stacking and may therefore decode with a somewhat lesser efficiency than other correct duplexes. Such inefficient base pairs are also indicated in Table 1. Our supplements to Crick's rules are supported experimentally (e.g., Munz et al., 1981; Osawa et al., 1992; Boren et al., 1993; Curran, 1995; Inagaki et al., 1995).

Besides providing the wobble rules of all unmodified residues (Lim & Venclovas, 1992; Lim, 1995), the wobble rules for modified residues a_1 have also been derived (Lim, 1994). The validity of these rules is also confirmed by all known modifications of a_1 (Table 1). Note that models based on the conformational characteristics of modified nucleotides and the assumption that A and C in a_1 can adopt protonated forms also allow explanation of the action of wobble nucleoside modifications on wobble base pairing (Yokoyama & Nishimura, 1995). However, nucleoside conformational characteristics are determined by weak intranucleotide interactions (~2 kJ/mol; Yokoyama et al., 1985) and are, therefore, not of crucial importance for discriminating the correct and wrong wobble pairs (for detailed analysis of a correlation between our and the Yokoyama-Nishimura models, see Curran, 1998).

The misreading rules

By our model, misreading frequencies should occur at levels of about $10^{-3.5\times M}$, where M is the number of

uncompensated lost hydrogen and ionic bonds in mispaired duplexes under the influence of the codon interribose bond SREs. Virtually all natural misreading events should involve duplexes with $M = 1$ because duplexes with $M > 1$ should give negligibly small levels of misreading ($\sim 10^{-7}$ and less). Therefore, to derive misreading rules, we identified all of the wrong duplexes that have an uncompensated loss of only one hydrogen or ionic bond.

Our analysis (for details, see Materials and Methods) shows that wrong duplexes with an uncompensated loss of only one hydrogen or ionic bond can contain only one "incorrect" base pair. At c_1 and c_2 such incorrect pairs are the mismatches AC, CA, GU, and UG. With regard to errors at c_3, because of the high mobility of a_1, virtually all pairs that are incompatible with the wobble rules (above) contain just a single lost hydrogen or ionic bond. For example, even I_1G_3 and G_1G_3 can be formed with a loss of only one bond when a_1 has *syn*-conformation. The only hard exception occurs for CC mispairs, which leads to two uncompensated, disrupted bonds (see Materials and Methods and Fig. 2C).

The derived codon misreading rules are given in Table 2. The rules are in complete accord with the available data on misreading in *Escherichia coli* (Parker, 1989). Those data show that (1) errors are observed at only one of the three codon positions; (2) errors at c_1 or c_2 are provided only by AC, CA, UG, and GU; (3) errors at c_3 are provided by a_1c_3 in which an uncompensated loss of one hydrogen bond is observed; and finally (4) the average error frequency is about $10^{-3.5}$. All of these observations correspond exactly with our results.

TABLE 2. Rules for codon misreading.

Rule 1. Misreading errors should occur at only one of the three codon positions.

Rule 2. Misreading errors at the first and second codon positions:

Nucleoside on the anticodon second or third position	Nucleoside misread on the codon
U	G
C	A
A	C
G	U

Rule 3. Misreading errors at the third codon position:

Nucleoside on the anticodon first position	Nucleoside misread on the codon
C	A, U
G	A, G
I	G
S^2U	U, C
Se^2U	U, C
Um	U, C
xm^5U	U, C
xo^5U	C
k^2C (lysidine)	U

However, it is important to note that the level of error can vary depending on other factors such as aminoacyl–tRNA concentrations. For example, Calderone et al. (1996) observe high levels of misincorporation of Lys for Arg at AGA codons, which is normally read by a rare arginine tRNA. That misreading event obeys our rules in that it involves a UG pair at a_2c_2, but it occurs at frequencies much greater than the typical misreading event. Another example occurs in animal mitochondria, in which specific codons actually lack tRNAs that our model would consider cognate (Tomita et al., 1999). In those cases the codons are read by tRNAs that will form duplexes with an uncompensated loss of only one bond, that is, the reading of the codon will occur in accord with our misreading rules. Semantically, of course, such reading is not "errant" in those systems. It would be of interest to determine whether such unusual decoding occurs with near-normal rates or efficiencies.

High-level codon misreading caused by A_2C_2, C_2A_2, and U_2G_2 and its prevention by the choice or modification of the first anticodon residue

Surprisingly, we revealed (for stereochemical details, see Materials and Methods) that in certain duplexes that have the sterically "soft" wobble pairs P_Y^P_Y (Fig. 2), the broken bond in mismatches A_2C_2, C_2A_2, and U_2G_2 can be compensated with alternate bonds, that is, such duplexes are not incorrect in our model. For duplexes with A_2C_2 or C_2A_2 mismatches to be correct, the duplexes must also have either A_3U_1 or U_3A_1 in addition to the wobble P_Y^P_Y pair. For duplexes with U_2G_2 to be correct, any canonical base pair in a_3c_1 will suffice. Soft P_Y^P_Y wobble pairs do not allow wrong duplexes with the asymmetric pair G_2U_2 to be assembled without uncompensated losses of hydrogen and ionic bonds.

Thus, the anticodons P_YAA, P_YAU, P_YCA, P_YCU, P_YUU P_YUC P_YUA, and P_YUG can form fundamentally correct duplexes containing mismatches A_2C_2, C_2A_2, and U_2G_2. If this were to occur during translation, these anticodons would cause very high frequency second-position errors. How are such errors avoided? It may be seen in Table 3 that the anticodons P_YAA, P_YAU, P_YCA, P_YCU, P_YUU P_YUC P_YUA, and P_YUG read codons from *all* eight mixed codon boxes and *none* from the unmixed boxes. In mixed boxes, P_Y at a_1 is always modified to prevent the reading of the P_Y-ending codons. Apparently, restricted wobbling in mixed boxes has the additional benefit of prohibiting high-level cross-box misreading due to noncanonical pairing A_2C_2, C_2A_2, and U_2G_2.

Unmodified U in a_1 recognizes U, C, A, and G in c_3 (Table 1). Consequently, the presence of this U in the anticodons of the codon mixed boxes should lead to both intra-box and cross-box misreading at levels that

TABLE 3. The genetic code.

UUU	Phe	**UCU**	**Ser**	UAU	Tyr	UGU	Cys
UUC	Phe	**UCC**	**Ser**	UAC	Tyr	UGC	Cys
UUA	Leu	**UCA**	**Ser**	UAA	Stop	UGA	Stop
UUG	Leu	**UCG**	**Ser**	UAG	Stop	UGG	Trp
CUU	**Leu**	**CCU**	**Pro**	CAU	His	**CGU**	**Arg**
CUC	**Leu**	**CCC**	**Pro**	CAC	His	**CGC**	**Arg**
CUA	**Leu**	**CCA**	**Pro**	CAA	Gln	**CGA**	**Arg**
CUG	**Leu**	**CCG**	**Pro**	CAG	Gln	**CGG**	**Arg**
AUU	Ile	**ACU**	**Thr**	AAU	Asn	AGU	Ser
AUC	Ile	**ACC**	**Thr**	AAC	Asn	AGC	Ser
AUA	Ile	**ACA**	**Thr**	AAA	Lys	AGA	Arg
AUG	Met	**ACG**	**Thr**	AAG	Lys	AGG	Arg
GUU	**Val**	**GCU**	**Ala**	GAU	Asp	**GGU**	**Gly**
GUC	**Val**	**GCC**	**Ala**	GAC	Asp	**GGC**	**Gly**
GUA	**Val**	**GCA**	**Ala**	GAA	Glu	**GGA**	**Gly**
GUG	**Val**	**GCG**	**Ala**	GAG	Glu	**GGG**	**Gly**

The eight codon boxes in which all the four codons specify the same amino acid are the unmixed boxes (bold letters). The other eight boxes are mixed.

should significantly exceed the experimentally observed level of $10^{-3.5}$. But it is important to note that the level of cross-box misreading should be less than that of intrabox misreading, as mismatches A_2C_2, C_2A_2, and U_2G_2 slightly distort the canonical orientation of the glycosyl bonds. These theoretical predictions can be subjected to direct experimental tests.

DISCUSSION

Strengths and limitations of the model and its further elaboration

The model is based on only three well-known characteristics of translation: mRNA codons are decoded by both tRNA anticodons and protein anticodons, the level of misreading is about $10^{-3.5}$, and the ribosomal cycle time is about $10^{-1.9}$ s. These characteristics impose strong restrictions on a choice of basic stereochemical and energetic parameters of decoding. Together, these restrictions provide a straightforward model by which ribosomes exploit the most fundamental and well-understood features of RNA structure to identify properly paired duplexes. The model can be refined with data, but is already powerful enough to explain all of the available data on decoding accuracy. Straightforward extensions of the model should allow development of models for other aspects of translation; frameshifting, hopping, and tmRNA translation, as examples. Further extensions may also facilitate understanding of the structures and functions of other RNAs, such as ribozymes.

One limitation of the model is that it does not account for influences by the tRNA anticodon arm residues outside of the anticodon, including modified nucleotides. Our model provides a first-order estimation of duplex correctness and considers only codon and anticodon nucleotides. Anticodon arm nucleotides outside of the anticodon contribute to the overall efficiency of decoding (see, e.g., Yarus, 1982). However, generally, the nucleotides and modifications of nucleotides outside of the anticodon have effects of less than 10-fold on translational efficiency (Yarus et al., 1986; Esberg & Björk, 1995; Li et al., 1997; Curran, 1998; Qiang et al., 1998) and duplex stability (Grosjean & Houssier, 1990). These effects may fine-tune decoding, but are much smaller than the $>$1,000-fold effects that our model addresses. However, improved understanding of the structural effects of modified nucleosides, for example, would permit a refinement of our model to account for these effects.

Nucleoside modifications *within* the anticodon are directly relevant to the current model. The roles of these bases on duplex stability have been addressed in a previous stereochemical study (Lim, 1994). In addition, the experimental data on the roles of modified nucleosides on codon recognition have been recently reviewed (Björk, 1992, 1998; Yokoyama & Nishimura, 1995; Curran, 1998). It is absolutely clear that modified nucleotides affect the decoding spectra of tRNAs, and all of those data are in complete agreement with the current model for determining duplex correctness. These roles of the various modified bases on extending and restricting the wobbling by modified nucleosides are summarized in Table 1.

One factor not accounted for by our model is the effects of interactions between the duplexes in the P- and A-sites on aminoacyl–tRNA selection. There is evidence for such interaction (Smith & Yarus, 1989; Curran, 1995), and they may account for much the "context effects" that have left imprints in gene structure (Yarus & Folley, 1985; Buckingham, 1990; Yarus & Curran, 1992). The mechanisms are not known, but like anticodon arm effects, these interduplex effects are of a much smaller magnitude than the large differences between correct and incorrect complexes that we address in our model. But, thanks to impressive progress in investigations of the structure of the ribosome, even these subtle effects may be incorporated into a refined stereochemical model of decoding in the near future.

Our model is compatible with other observations on tRNA-ribosomal association and with both proofreading and allosteric interaction kinetic models for aminoacyl–tRNA selection

Our model proposes that duplex recognition will occur in stages: duplex formation followed by the application of the SREs of the interribose bonds. Others have observed that tRNA–ribosome interaction occurs in stages

(Lake, 1977; Robbins & Hardesty, 1983; Moazed & Noller, 1989). Although the codon recognition steps addressed here need not correspond directly to the large-scale conformational steps observed previously, our analysis provides a robust confirmation that codon recognition must occur in stages.

Our model focuses only on the mechanism for duplex recognition and does not depend on any particular ribosomal kinetic scheme. Ribosomes must recognize correct duplexes in the A-site regardless of how the ternary complex and aminoacyl–tRNA components are processed. Indeed, one of our motivations for developing the model is the observations by Thompson & Karim (1982) that cognate duplexes can be effectively recognized on ribosomes that do not have tRNA in the E-site and do not hydrolyze GTP in a proofreading step. Our model is, however, consistent with either type of kinetic scheme. In proofreading models, GTP hydrolysis could separate recognition of the correct duplex into multiple steps to increase ribosomal rate without loss of accuracy as suggested by Thompson & Karim (1982), or recognition of the correct duplex could trigger GTP hydrolysis (Pape et al., 1998). Similarly, in an allosteric three-site model, it could be the recognition of the stable, correct duplex that allows for an allosteric change expelling the deacylated tRNA from the E-site (Nierhaus, 1993).

Candidates for SREs of the codon interribose bonds

Theoretically, SREs could be components of the ribosome or could be the wobble pair of the P-site duplex. Previously it was shown (Lim, 1997; Lim & Aglyamova, 1998) that the P-site wobble pair could serve as SRE regardless of the mutual orientation of the P- and A-site tRNAs. Moreover, the crystal structure of 70S ribosome-containing tRNAs (Cate et al., 1999) and the crystal structure of the 30S ribosomal subunit (Carter et al., 2000) show that the ribosomal A site is large enough to allow duplexes to form away from the P-site wobble pair, followed by their succeeding movement into the vicinity of this pair. Thus, the duplex could form outside of the influence of this putative SRE and then move to it, as is required by our model.

Alternatively, ribosomal components could serve as SREs. The data of Yoshizawa et al. (1999) strongly suggest that the N1 atoms of the universally conserved adenines A1492 and A1493 of 16S rRNA contact two (OH)2′ groups of the A-site codon. Yoshizawa et al. (1999) proposed that A1492 and A1493 form hydrogen bonds N1•••(OH)2′ to help fix the codon A-form. Other schemes of hydrogen bonding between the codon and adenines 1492(3) have also been proposed (VanLoock et al., 1999; Carter et al., 2000). In many variants of an interaction with the duplexes, two adenines cannot fix the A-form conformation of even two of the three codon residues. However, they could be used as SREs for the interribose bonds (Fig. 6), which would fix the A-form. As described above and outlined in Figure 6, this could occur in either of two ways. In these ways, the two adenines 1492(3) could fix all three codon residues in the A-form conformation.

There are X-ray data that, at first glance, are in conflict with the results of Yoshizawa et al. (1999). In the crystal structure of 70S ribosome-containing tRNAs (Cate et al., 1999), the formed A-site duplex is more than 15 Å from the N1 positions of A1492 and A1493. This "discrepancy" correlates well with our model if A1492 and A1493 are SREs of the interribose bonds. According to our model the crystal structures of 70S ribosome-containing tRNAs demonstrates the first stage of the duplex selection where duplexes assemble beyond the action of the steric restriction elements. The interaction between the codon and adenines 1492(3) detected by Yoshizawa et al. (1999) corresponds to the second stage where SREs (A1492 and A1493) and the duplex (its codon moiety) are drawn together. Recent structural data suggests that the A1492 and A1493 may move towards the codon. The crystal structure of the 30S ribosomal subunit (Carter et al., 2000) shows that the electron density for A1492 and A1493 is not consistent with a single conformation for these residues. Moreover, a conformational switch occurs (Pape et al., 2000) in the decoding region of 16S rRNA during aminoacyl–tRNA selection on the ribosome.

The above data demonstrate that the P-site wobble pair and A1492 and A1493 are likely candidates for SREs for the interribose bonds. Therefore, it will be of interest to determine the locations of the P-site wobble pair, and A1492 and A1493 in ribosomal complexes corresponding to the postacceptance of an aminoacyl–tRNA stage.

Wobble pairing, the structure of the mixed boxes, and their distribution in the genetic code dictionary

P_UP_U pairs at a_1c_3 are rare, but they do occur during translation. In such pairs, due to fixation of the codon in the A-form, a_1 carries out a large shift in the anticodon backbone relative to that needed for canonical pairs (Lim & Venclovas, 1992). This large shift of a_1 leads to a sterically strained stretching of the backbone section between the two first anticodon bases. Therefore, any wobble pair P_UP_U should be weakly effective. This stereochemical conclusion is strongly supported by experimental data on efficiency of the wobble pair I_1A_3. Munz et al. (1981) and Curran (1995) have shown that this pair is relatively ineffective.

Besides low efficiency of the wobble pairs P_UP_U, we have shown that to avoid mismatches at c_2, the anticodons of the mixed boxes should not form the wobble pairs P_YP_Y. Consequently, the decoding at c_3 should

basically be accomplished with P_YP_U and P_UP_Y, that is, the anticodons of the mixed boxes should recognize, as a rule, only the codons ending with A and/or G when a_1 is P_Y or the codons ending with U and/or C when a_1 is P_U. This correlates well with the structure of the mixed boxes (Table 3) where the codons ending with U and C specify, as a rule, one amino acid and the codon ending with A and G specify the other one.

Prohibition of mismatches at c_2 has also obvious implications for the distribution of the mixed boxes. Assuming that primitive codes specified fewer amino acids, then it is likely that mixed boxes originally specified single amino acids. However, under pressure to prevent errors at c_2, restricted a_1 wobbling may have functionally split the boxes. Then under pressure for an expanded code, one half of each box acquired a new meaning, creating mixed boxes. Conversely, without pressure to limit errors at the second position, a_1 wobble was not restricted in full boxes, which therefore did not have the opportunity to develop into mixed boxes.

MATERIALS AND METHODS

Calculations of the probability of incorrect reading and the average time required to overcome energetic barriers

At equilibrium, the probability of incorrect reading is $\exp(-\Delta G/RT)$. The in vivo observed level of misreading is about $10^{-3.5}$. At $T = 300$ K, $\exp(-\Delta G/RT)$ is equal to $10^{-3.5}$ when $\Delta G \approx 20$ kJ/mol. This means that at equilibrium, a difference of 20 kJ/mol between the free energy of wrong and right association is required to provide the observed level of misreading.

The average time (τ_0) required to overcome a barrier with the height H is equal to $h/kT \times \exp(H/RT)$. At $T = 300$ K, the value of h/kT is about 10^{-13} s. At $H = 3 \times 20$ kJ/mol and 4×20 kJ/mol, $\tau_0 = 10^{-2.5}$ s and $\sim$10 s, respectively. For background on the formulas, see Tinoco et al. (1995).

Central tenets

Analysis of the codon:anticodon duplexes was made using computer graphics methods. A complex was disallowed when at least one interatomic distance was less than the extreme limit (rarely observed steric overlaps; Ramachandran & Sasisekharan, 1968). A hydrogen bond D–H•••A was considered disrupted when HA exceeded the extreme limit distance between H and A and/or the angle DHA was less than 150°. Anions were formally replaced by oxygen (O^{2-} and F^- are the smallest anions with a radius of 1.4 Å; Pauling & Pauling, 1975). For interactions with hydrogen bond donors, anions were considered as hydrogen bond acceptors. Cation-ligand bonds were considered when duplex hydrogen bond acceptors could not form hydrogen bonds with water molecules for steric reasons or could not form the bonds without uncompensated losses of other bonds. Cations with a radius of 1 Å and tetrahedrally oriented bonds were used.

Nonstandard propeller twist and *syn*-conformation of bases in the duplex base pairs

The interribose bonds prohibit the changes in the mutual orientations of adjacent codon bases that are required for a duplex base pair to form a standard propeller twist. Therefore, we used the nonstandard propeller twist resulting from a positive rotation around the glycosyl bond (rotational angle χ) of the anticodon bases (Fig. 1A). The other ways of propeller twist formation are either sterically prohibited or require unacceptable shifts of the sugar–phosphate moiety of the anticodon residues. The nonstandard twist was first proposed by Lim (1994), and then a UU pair with the large propeller twist of this type was found in an RNA double helix (Baeyens et al., 1995).

In principle, the large nonstandard propeller twist can eliminate uncompensated losses of hydrogen and ionic bonds within base pairs (Fig. 1A). However, almost never can bond losses within the base pairs a_2c_2 and a_3c_1 be eliminated with the high nonstandard propeller twist (there are only several exceptions; see below). This is because the large positive rotation around the glycosyl bond in a_2 and a_3 is sterically prohibited, because a_2 and a_3 are sandwiched between a_1c_3 and a_3c_1 and between a_2c_2 and tRNA conserved purine base 37, respectively. Only the anticodon base a_1 can form the large nonstandard propeller twist. Previously the nonstandard twist was used to find allowed wobble base pairs, including pairs with a wide diversity of modifications of the wobble anticodon residue (Lim, 1994, 1995).

Increasing χ is inversely proportional to the distance from the axis of χ to the atom that should be shifted to avoid a bond loss. When the required increase for χ is very large ($\sim$30–40°), some base–base hydrogen bonds in a twisted base pair are significantly distorted. They can be restored by shifts of 1–2 Å of peripheral part of the anticodon base toward the anticodon 3′ end.

Another important application of the nonstandard twist is an avoidance of the simultaneous breakage of three hydrogen bonds during the disruption/formation of the GC pairs in the duplexes. For example, three base–base hydrogen bonds can be disrupted and replaced by base–solvent bonds in steps of one and then two bonds. One bond is replaced with the twist and the other two by removal a codon base from the minihelix.

Besides the nonstandard twist, we considered the *syn*-conformation of bases in the search for wrong duplexes that contain an uncompensated loss of only one bond (see below). (Note that the *syn*-conformation is not admissible in *correct* duplexes because it causes uncompensated losses of bonds due to steric restrictions of the duplex sugar–phosphate moiety.)

Pyrimidine:pyrimidine pairs P_Y^P_Y containing a base–water–base bridge

Because of the short distance between the glycosyl bonds, the P_YP_Y pairs UU, UC, and CC having two base–base hydrogen bonds cannot be incorporated into duplexes with the fixed codon A-form. Even the wobble pair a_1c_3, in which a_1 is mobile, cannot be formed of the "short" P_YP_Y without disallowed shifts of the anticodon sugar–phosphate moiety (Lim &

Venclovas, 1992). However, the pairs U^C and U^U in which one base–base hydrogen bond is replaced by a base–water–base bridge (Fig. 2A,B) have been observed (Holbrook et al., 1991; Cruse et al., 1994; Wang et al., 1996). The pair C^C is sterically prohibited (Fig. 2C). The distance between the glycosyl bonds in U^C and U^U is close to that in canonical base pairs, but the mutual orientation of the glycosyl bonds, especially in U^U, significantly differs from canonical one. Thus they form only wobble-type pairs. A water bridge is located in the inner part of the pair, especially in U^C, and its two bonds are pointed toward adjacent pairs and base 37 (Fig. 2A,B). These structural characteristics and the prohibition of the large nonstandard twist do not permit a_2c_2 and a_3c_1 of U^C and U^U to form without losses of two bonds.

Base pairs causing only one uncompensated loss of hydrogen or ionic bonds when the duplex codon is fixed in the A-form

The allowed wobble pairs a_1c_3 (Table 1) do not have uncompensated losses of the bonds. Hence, the pairs a_1c_3 with lost bonds are CU, CA, CC, GA, GG, IG, S^2UU, S^2UC, Se^2UU, Se^2UC, UmU, UmC, xm^5UU, xm^5UC, xo^5UC, k^2CU, k^2CC, and k^2CG. These pairs can be formed so that they will lose only one bond. This bond is either the c_2-c_3 interribose bond, a bond within the base pair, or a bond that is lost in the *syn*-conformation of bases. The exception is CC. The pair C^C is prohibited (Fig. 2C). Base–base hydrogen bonding in short pairs CC is the same as that shown in Figure 1A if one considers the six-membered ring of adenosine an analog of cytosine. A short CC can be formed after disruption of the c_2-c_3 interribose bond. When the "cytosine" in Figure 1A is the anticodon base and its atom N3 is protonated, the short CC has two base–base hydrogen bonds. The free polar atom N3 of C and/or NH_2 group(s) in the wobble short pair CC are shifted from the edge of a_2c_2 to its inner part. For this reason, a_2c_2 is an SRE for solvent molecules bonded to the N3 and NH_2. The action of this SRE cannot be eliminated by the large nonstandard or standard twist (after disruption of the c_2-c_3 interribose bond the standard twist can also be used). Thus, the short wobble pair CC loses two or more bonds, and one of them is the c_2-c_3 interribose bond.

The wrong duplexes with only one lost bond are observed when a_2c_2 or a_3c_1 is any one of the pairs AC, CA, GU, and UG. These P_UP_Y mismatches have quasi-canonical orientations of the glycosyl bonds (Fig. 3) that only slightly change the helix A-form. In the pair AC (Fig. 1B) N3 of C cannot form a bond with solvent because of a steric restriction created by H2 of A. For the alternative GU-like pairing (Fig. 1A) a loss of two bonds occurs. Protonation of N1 of A leads to a loss of a single bond.

In RNA double helices, the ribose (OH)2′ group of U in the pair GU participates in the network connecting (OH)2′(U) to NH_2(G) through an intermediate water molecule (e.g., Cruse et al., 1994). In oligodeoxynucleotides, an analogous bridge is formed between O2 and NH_2 of the same residues. One water molecule can also simultaneously form both bridges without strong distortions of hydrogen bonds (Fig. 3A). In the fixed codon A-form, the pairs GU and UG are prohibited in c_1 and c_2. Regardless of the number of bonds (two or three) formed by a bridging water molecule with the pair GU, one of its bond cannot interact with solvent because of steric restrictions created by edges of adjacent base pairs including conserved purine base 37. A loss of this bond can be eliminated by only simultaneous shifts of about 1–2 Å in both glycosyl bonds of the GU pair from the position in the immediate vicinity of the glycosyl bonds of the canonical pair. But such shifts lead to the disruption of the codon interribose bonds.

Wobbling A_2C_2, C_2A_2, and U_2G_2 in the pair a_2c_2

As discussed above, steric restrictions created by adjacent base pairs and base 37 leads to the loss of a single bond in the pairs AC, CA, GU, and UG located in c_1 or c_2. However, we have found duplexes in which these restrictions are absent, but these duplexes cannot be formed by the anticodon sets that are used in vivo. In the presence of the allowed wobble pairs U_1^U_3 and U_1^C_3, the pairs A_2C_2, C_2A_2, and U_2G_2 can exist without uncompensated losses of hydrogen and ionic bonds. A flexible bridging water molecule located in the central part of P_Y^P_Y (Fig. 2) permits the large positive change of the angle χ in a_2 to form the large nonstandard propeller twist in A_2C_2 and C_2A_2 (Fig. 3B) that permits avoidance of a loss of the bonds in these pairs. In the case of U_2G_2, a flexible bridging water molecule permits avoidance of a loss of the bond by a water molecule hydrogen bonded to the guanine NH bond (Fig. 3A). The presence of the other wobble pairs a_1c_3 sterically prohibits the large positive change of χ in a_2 and together with a_3c_1 prohibits one bond of a water molecule bonded to the guanine NH bond.

To provide incorporation of AC and CA into c_2, besides the wobble pairs P_Y^P_Y, position c_1 should be occupied by AU or UA. These pair are sterically more soft than GC and CG and they do not have the NH_2 group in the minor groove of the minihelx. These NH_2 groups in G_3C_1 and C_3G_1 are located approximately at the same place in the minor groove and strongly counteract the interaction of N3 of C with solvent in the pair A_2C_2. The pair C_2A_2 also requires the presence of AU or UA in c_1. In C_2A_2, the twist should be very large (~30–40°), greater than that in A_2C_2 because cytosine polar atom N3 that should be shifted is located close to the axis of χ. Formation of the required twist in C_2A_2 is accompanied by disruption of the base–base hydrogen bond in this pair. Therefore, to restore this bond, a shift of the peripheral part of C in C_2A_2 toward a_3c_1 should occur. Such a shift can occur only when position c_1 is occupied by AU or UA, which are sterically softer than GC and CG.

As to GU and UG in c_2, the pair U_2G_2 is formed in the presence of the allowed wobble pairs P_Y^P_Y regardless of the type of canonical pair in c_1. The asymmetric pair G_2U_2 cannot be used because a water molecule hydrogen bonded to NH_2 of G is far removed from the bridging water in the wobble pairs P_Y^P_Y and cannot interact with (OH)2′ of U without disruption of the codon interribose bond.

ACKNOWLEDGMENTS

We thank M.B. Garber and K.H. Nierhaus for constructive comments. This work was supported by National Institutes of Health Grants GM 52643 and GM 58425 and a Wake Forest

University, Computer-Enhanced Learning Initiative grant (J.F.C.) and by Grant 99–04–48653 from the Russian Foundation for Basic Research and Grant INTAS 96–1706 (V.I.L.).

Received October 19, 2000; returned for revision November 14, 2000; revised manuscript received April 26, 2001

NOTE ADDED IN PROOF

While this manuscript was in press, V. Ramakrishnan and co-workers published crystal structures of the 30S ribosomal subunit containing mRNA paired to an anticodon arm mimic in the A site (Ogle JM, Brodersen DE, Clemons WM Jr, Tarry MJ, Carter AP, & Ramakrishnan V. 2001. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science 292*:897–902). In this complex the minor groove of the codon: anticodon duplex interacts with the universally conserved residues A1492, A1493, and G530 of 16S rRNA. Note that bases A1492, A1493, and G530 can rotate about their glycosyl bonds. Such rotation means two things relevant to our model. First, hydrogen bonds between these bases and the duplex may form and break independently of bonds between the rRNA backbone and the duplex. This independence allows the complex to form and dissociate without violating rule $N + M < 4$. Second, hydrogen bonds between these bases and the duplex cannot ensure accuracy. When noncognate duplexes occupy the A site, these bases can simply rotate to find alternate pairing partners with solvent molecules. However, these rRNA residues ensure accuracy in another way: They are SREs of the codon inter-ribose bonds. In the presence of these SREs, noncognate duplexes are not stable because their disrupted inter-ribose bonds cannot find alternative pairing partners.

REFERENCES

Agris PF. 1991. Wobble position modified nucleosides evolved to select transfer RNA codon recognition: A modified-wobble hypothesis. *Biochimie 73*:1345–1349.

Baeyens KJ, De Bondt HL, Holbrook SR. 1995. Structure of an RNA double helix including UU pairs in an internal loop. *Nat Struct Biol 2*:56–62.

Björk GR. 1992. The role of modified nucleosides in tRNA interactions. In: Hatfield DL, Lee BJ, Pirtle RM, eds. *Transfer RNA in protein synthesis*. Boca Raton, FL: CRC Press. pp 23–85.

Björk GR. 1998. Modified nucleosides at positions 34 and 37 of tRNAs and their predicted coding capacities. In: Grosjean H, Benne R, eds. *Modification and editing of RNA*. Washington, DC: ASM Press. pp 577–581.

Boren T, Elias P, Samuelsson T, Claesson C, Barciszewska M, Gehrke CW, Kuo KC, Lustig F. 1993. Undiscriminating codon reading with adenosine in the wobble position. *J Mol Biol 230*:739–749.

Buckingham RH. 1990. Codon context. *Experientia 46*:1126–1133.

Burkhardt N, Jünemann R, Spahn CMT, Nierhaus KH. 1998. Ribosomal tRNA binding sites: Three-site models of translation. *Crit Rev Biochem Mol Biol 33*:95–149.

Calderone TL, Stevens RD, Oas TG. 1996. High-level misincorporation of Lys for Arg at AGA codons in a fusion protein expressed in *E. coli*. *J Mol Biol 262*:407–412.

Carter AP, Clemons WM, Broderson DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V. 2000. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature 407*:340–348.

Cate JH, Yusupov MM, Yusupova GZ, Earnest TN, Noller HF. 1999. X-ray structures of 70S ribosome functional complexes. *Science 285*:2095–2104.

Conn GL, Draper DE, Lattman EE, Gittis AG. 1999. Crystal structure of a conserved ribosomal protein–RNA complex. *Science 284*: 1171–1174.

Crick FHC. 1966. Codon-anticodon pairing: The wobble hypothesis. *J Mol Biol 19*:548–555.

Cruse WBT, Saludjan P, Biala E, Strazewski P, Prange T, Kennard O. 1994. Structure of a mispaired RNA double helix at 1.6-Å resolution and implications for the prediction of RNA secondary structure. *Proc Natl Acad Sci USA 91*:4160–4164.

Curran JF. 1995. Decoding with the A:I wobble pair is inefficient. *Nucleic Acids Res 23*:683–688.

Curran JF. 1998. Modified nucleosides in translation. In: Grosjean H, Benne R, eds. *Modification and editing of RNA*. Washington, DC: ASM Press. pp 577–581.

Dirheimer G, Keith G, Dumas P, Westhof E. 1995. Primary, secondary, and tertiary structure of tRNAs. In: Söll D, RajBhandary U, eds. *tRNA: Structure, biosynthesis, and function.* Washington, DC: ASM Press. pp 93–126.

Esberg B, Björk GR. 1995. The methylthio group (ms2) of N6-(4-hydroxyisopentenyl)-2-methylthioadenosine (ms2io6A) present next to the anticodon contributes to the decoding efficiency of the tRNA. *J Bacteriol 177*:1967–1975.

Grosjean H, Houssier C. 1990. Codon recognition: Evaluation of the effects of modified bases in the anticodon loop of tRNA using the temperature-jump relaxation method. *J Chromatogr Library 45A*: A255–A295.

Grosjean HJ, de Henan S, Crothers DM. 1978. On the physical basis for ambiguity in genetic coding interactions. *Proc Natl Acad Sci USA 75*:610–614.

Gurskaya GV. 1968. *The molecular structure of amino acids determination by X-ray diffraction analysis*. New York, NY: Consultants Bureau.

Holbrook SR, Cheong C, Tinoco I, Kim S-H. 1991. Crystal structure of an RNA double helix incorporating a track of non-Watson–Crick base-pairs. *Nature (London) 353*:579–581.

Hopfield JJ. 1974. Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci USA 71*:4135–4139.

Inagaki Y, Kojima A, Bessho Y, Hori H, Ohama T, Osawa S. 1995. Translation of synonymous codons in family boxes by *Mycoplasma capricolum* tRNAs with unmodified uridine or adenosine at the first anticodon position. *J Mol Biol 251*:486–492.

Jeffrey GA, Maluszynska H, Mitra J. 1985. Hydrogen bonding in nucleosides and nucleotides. *Int J Biol Macromol 7*:336–348.

Kurland CG, Jörgensen F, Richter A, Ehrenberg M, Bilgin N, Rojas AM. 1990. Through the accuracy window. In: Hill WE, Dahlberg A, Garrett RA, Moore PB, Schlessinger D, Warner JR, eds. *The ribosome: Structure, function and evolution*. Washington, DC: ASM Press. pp 513–526.

Lake JA. 1977. Aminoacyl–tRNA binding at the recognition site is the first step in the elongation cycle of protein synthesis. *Proc Natl Acad Sci USA 11*:4248–4251.

Li J-n, Esberg B, Curran JF, Björk, GR. 1997. Three modified nucleosides present in the anticodon stem and loop influence the in vivo aa-tRNA selection in a tRNA-dependent manner. *J Mol Biol 271*:209–221.

Lim VI. 1994. Analysis of action of wobble nucleoside modifications on codon-anticodon pairing within the ribosome. *J Mol Biol 240*: 8–19.

Lim VI. 1995. Analysis of action of the wobble adenine on codon reading within the ribosome. *J Mol Biol 252*:277–282.

Lim VI. 1997. Analysis of interactions between the codon-anticodon duplexes within the ribosome: Their role in translation. *J Mol Biol 266*:877–890.

Lim VI, Aglyamova GV. 1998. Mutual orientation of tRNAs and interactions between the codon-anticodon duplexes within the ribosome: A stereochemical analysis. *Biol Chem 379*:773–781.

Lim VI, Venclovas C. 1992. Codon-anticodon pairing. A model for interacting codon-anticodon duplexes located at the ribosomal A- and P-sites. *FEBS Letters 313*:133–137.

Moazed D, Noller HF. 1989. Intermediate states in the movement of transfer RNA in the ribosome. *Nature 342*:142–148.

Moras D, Comarmond MB, Fischer J, Weiss R, Thierry JC, Ebel JP,

Giege R. 1980. Crystal structure of yeast $tRNA^{Asp}$. *Nature 288*: 669–674.
Munz P, Leupold U, Agris P, Kohli J. 1981. In vivo decoding rules in *Schizosaccharomyces pombe* are at variance with in vitro data. *Nature 294*:187–188.
Nierhaus KH. 1993. Solution of the ribosome riddle: How the ribosome selects the correct aminoacyl–tRNA out of 41 similar contestants. *Mol Microbiol 9*:661–669.
Ninio J. 1975. Kinetic amplification of enzyme discrimination. *Biochimie 57*:587–595.
Osawa S, Jukes TH, Watanabe K, Muto A. 1992. Recent evidence for evolution of the genetic code. *Microbiol Rev 56*:229–264.
Pape T, Wintermeyer W, Rodnina MV. 1998. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl–tRNA to the A site of the ribosome. *EMBO J 17*:7490–7497.
Pape T, Wintermeyer W, Rodnina MV. 2000. Conformational switch in the decoding region of 16S rRNA during aminoacyl–tRNA selection on the ribosome. *Nat Struct Biol 7*:104–107.
Parker J. 1989. Errors and alternatives in reading the universal genetic code. *Microbiol Rev 53*:273–298.
Pauling L, Pauling P. 1975. *Chemistry*. San-Francisco: WH Freeman & Company.
Potapov AP, Triana-Alonso FG, Nierhaus KH. 1995. Ribosomal decoding processes at codons in the A or P sites depend differently on 2′-OH groups. *J Biol Chem 270*:17680–17684.
Qiang Q, Curran JF, Björk GR. 1998. The methyl group of the N6-methyl-N6-threonylcarbamoyladenosine in tRNA of *Escherichia coli* modestly improves the efficiency of the tRNA. *J Bacteriol 180*:1808–1813.
Ramachandran GN, Sasisekharan V. 1968. Conformation of polypeptides and proteins. *Advan Protein Chem 28*:283–437.
Robbins D, Hardesty B. 1983. Comparison of ribosomal entry and acceptor transfer RNA binding sites on *E. coli* 70 S ribosomes: Fluorescence energy transfer measurements from $Phe\text{-}tRNA^{Phe}$ to the 3′ end of ribosomal RNA. *Biochemistry 22*:5675–5679.
Saenger W. 1984. *Principles of nucleic acid structure*. New York, NY: Springer-Verlag.
Smith D, Yarus M. 1989. tRNA-tRNA interactions within cellular ribosomes. *Proc Natl Acad Sci USA 86*:4397–4401.
Song H, Mugnier P, Das AK, Webb HM, Evans DR, Tuite MF, Hemmings BA, Barford D. 2000. The crystal structure of human eukaryotic release factor eRF1-Mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell 100*:311–321.
Thompson RC, Karim AM. 1982. The accuracy of protein biosynthesis is limited by its speed: High fidelity selection by ribosomes of aa-tRNA ternary complexes containing GTP[γS]. *Proc Natl Acad Sci USA 79*:4922–4926.
Tinoco I Jr, Sauer K, Wang JC. 1995. *Physical chemistry: Principles and applications in the biological sciences.* Englewood Cliffs, NJ: Prentice Hall.
Tomita K, Ueda T, Ishiwa S, Crain PF, McCloskey JA, Watanabe K. 1999. Codon reading patterns in *Drosophila melanogaster* mitochondria based on their tRNA sequences: A unique wobble rule in animal mitochondria. *Nucleic Acids Res 27*:4291–4297.
Ts'o POP. 1974. Bases, nucleosides, and nucleotides. In: Ts'o POP, ed. *Basic principles in nucleic acid chemistry.* New York, NY: Academic Press. pp 453–584.
VanLoock MS, Easterwood TR, Harvey SC. 1999. Major groove binding of the tRNA/mRNA complex to the 16 S ribosomal RNA decoding site. *J Mol Biol 285*:2069–2078.
Wang Y-X, Huang S, Draper DE. 1996. Structure of a UU pair within a conserved ribosomal RNA hairpin. *Nucleic Acids Res 24*:2666–2672.
Wimberly BT, Guymon R, McCutcheon JP, White SW, Ramakrishnan V. 1999. A detailed view of a ribosomal active site: The structure of the L11-RNA complex. *Cell 97*:491–502.
Yarus M. 1982. Translational efficiencies of tRNAs: Uses of an extended anticodon. *Science 218*:646–652.
Yarus M. 1992. Proofreading, NTPases and translation: Constraints on accurate biochemistry. *Trends Biochem Sci 17*:130–133.
Yarus M, Cline SW, Wier P, Breeden L, Thompson RC. 1986. Actions of the anticodon arm in translation on the phenotypes of RNA mutants. *J Mol Biol 192*:235–255.
Yarus M, Curran JF. 1992. The translational context effect. In: Hatfield DL, Lee BJ, Pirtle RM, eds. *Transfer RNA in protein synthesis*. Boca Raton, FL: CRC Press. pp 319–365.
Yarus M, Folley LS. 1985. Sense codons are found in specific contexts. *J Mol Biol 182*:529–540.
Yokoyama S, Nishimura S. 1995. Modified nucleoside and codon recognition. In: Söll D, RajBhandary UL, eds. *tRNA: Structure, biosynthesis, and function*. Washington, DC: ASM Press. pp 207–223.
Yokoyama S, Watanabe K, Murao H, Ishikura Z, Yamaizumi S, Nishimura S, Miyazawa T. 1985. Molecular mechanism of codon recognition by tRNA species with modified uridine in the first position of the anticodon. *Proc Natl Acad Sci USA 82*:4905–4909.
Yoshizawa S, Fourmy D, Puglisi JD. 1999. Recognition of the codon–anticodon helix by ribosomal RNA. *Science 285*:1722–1725.