Surprising contribution to aminoacylation and translation of non-Watson–Crick pairs in tRNA

William H. McClain*

Department of Bacteriology, University of Wisconsin, 420 Henry Mall, Madison, WI 53706-1569

Communicated by Sidney Altman, Yale University, New Haven, CT, January 24, 2006 (received for review November 25, 2005)

Molecules of transfer RNA (tRNA) typically contain four stems composed of Watson–Crick (W-C) base pairs and infrequent mispairs such as G-U and A-C. The latter mispairs are fundamental units of RNA secondary structure found in nearly every class of RNA and are nearly isomorphous to W-C pairs. Therefore, they often substitute for G-C or A-U base pairs. The mispairs also have unique chemical, structural, and dynamic conformational properties, which can only be partially mimicked by W-C base pairs. Here, I characterize the identities and tasks of six mutant G-U and A-C mispairs in Escherichia coli tRNA^{Gly} using genetic and bioinformatic tools and show that mispairs are significantly more important for aminoacylation and translation than previously realized. Mispairs boost aminoacylation and translation primarily because they activate tRNA by means of their conformational flexibility. The statistical preservation of the six mutant mispair sites across tRNA^{Gly} in many organisms points to a fundamental structure–function signature within tRNA^{Gly} with possible analogous missions in other RNAs.

Results and Discussion

Rationale. I am unaware of an aminoacylation system that contains a functionally significant and conserved G-U mispair similar to that in tRNA^{Ala}. However, several nonconserved and variously located G-U pairs in wild-type (WT) and mutant tRNAs can contribute to charging in the aspartic acid and alanine systems (6–8). Given the latter observation, I hypothesize that any one member of a group of mispairs at various positions may promote tRNA function.

I addressed the above hypothesis by first looking for mispairs common to tRNAs and then using a structure–function analysis to examine the importance of a common mispair to tRNA function. To this end, I identified mispairs in the four helical stems of tRNA in 1,550 genomic tRNA sequences from cytoplasmic sources that are nonredundant and derived from organisms of all three phylogenetic domains. The analysis revealed that G-U is the most frequent mispair, and although G-U is numerous at the base of the T stem (position 49–65), it and other mispairs distribute prominently over much of the tRNA sequence (Fig. 1). I studied the 49–65 G-U pair by mutating it to G-C in a nonsense suppressor tRNA system. I reasoned that if the G-C substitution functionally impaired the tRNA, then the isolation and characterization of mutant strains with distinctive second-site compensating mutations could identify a set of mispairs that compensated for the mutant G-C W-C pair and promoted tRNA function. Finally, in a recursive step, I determined the statistical occurrence of mispairs at compensating sites in the corresponding genomic tRNA sequences. A significant correlation between mutationally observed mispair sites and nature’s genomic mispair sites would provide further support for the hypothesis.

Conversion of G-U to G-C Inactivates tRNA Function. I chose the isoaacceptor of Escherichia coli tRNA^{Gly} (GCC anticodon) (Fig. 2) for analysis because its acceptor specificity is well studied (9, 10). I changed G-U at position 49–65 to G-C by synthesizing a gene for opal suppressor tRNA^{Gly} in plasmid pGFIB so the transcribed tRNA contained a UCA anticodon, which is complementary to the UGA opal triplet in mRNA. To avoid noncognate aminoacyl-tRNA synthetase competition for tRNA^{Gly}, I used the opal rather than the amber suppression system because the amber tRNA inserts a mixture of amino acids (9). Two versions of the gene, one with G-U at 49–65 (native opal suppressor tRNA^{Gly}) and the other with a G-C substitution at this position (opal suppressor C65tRNA^{Gly}), constitutively produce tRNA in 2.1-fold excess relative to the comparable chromosomal tRNA^{Gly} (11). I separately transformed the two plasmids into cells containing lac opal mutation U4 in a lacI-Z fusion system such that cells producing an active opal suppressor tRNA form blue colonies on 5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside (X-Gal) indicator plates, whereas cells lacking a suppressor tRNA, or expressing an inactive suppressor tRNA, form white colonies (9, 12). When tested, native opal suppressor tRNA^{Gly} formed blue colonies on X-Gal plates, whereas opal suppressor C65tRNA^{Gly} formed white to light blue colonies. A Northern blot analysis demonstrated that opal suppressor...
tRNA<sup>gly</sup> and C65tRNA<sup>gly</sup> are present in cells in similar amounts (Fig. 3 and Table 1). Thus, changing this G-U mispair to a G-C base pair damages one or more functions of the molecule.

**Activation of Damaged tRNA.** I sought gain-of-function derivatives of suppressor C65tRNA<sup>gly</sup> by hydroxylamine mutagenesis of plasmid pGFIB carrying this tRNA gene, transforming the treated plasmid into XAC/H20862<sup>U4</sup> cells, and screening for blue colonies on X-Gal indicator plates. Hydroxylamine is quite effective on tRNA genes (13), generating G-U or A-C mispairs in stem segments. I found 20 blue colonies among ~11,000 colonies examined. I recloned the 20 tRNA genes into fresh pGFIB plasmid and by DNA sequencing found that 17 of 20 isolates were double mutants containing the original C65 and a secondary base substitution. I named the double mutants C65XY, where X designates the mutant base substitution in the transcribed tRNA and Y its cloverleaf position number. Fig. 2 shows the locations and nature of the mispairs in 17 of the tRNAs. The mispairs do not segregate to one face of the tRNA, but rather they encompass the molecule when viewed on the crystal structure of yeast tRNA<sup>Pho</sup>. One of the three remaining isolates reverted to the original U65 residue, another contained no change relative to starting C65tRNA<sup>gly</sup>, and another contained too little DNA to generate a sequence. The mutagenesis probably was not saturating because I observed only one isolate at each of three sites.

The structural density of the mutations led me to construct derivative genes in native opal suppressor tRNA<sup>gly</sup> such that each new gene contained just a compensating mutation. I determined the in vivo aminoacylation specificity of every tRNA<sup>gly</sup> mutant by sequencing the N-terminal portion of a reporter dihydrofolate reductase (DHFR) protein carrying an opal triplet at codon 3 in its mRNA. All C65 double and U65 single mutants inserted only glycine in the reporter protein (Table 1). Native opal suppressor tRNA<sup>gly</sup> was also specific for glycine, whereas in three independent

---

**Fig. 1.** Distribution of G-U and other mispairs in 1,550 tRNA nonredundant sequences representing 20 accepting types.

**Fig. 2.** Cloverleaf structure of *E. coli* tRNA<sup>gly</sup> (GCC anticodon isoacceptor). The base numbers in the tRNA reflect the standard nomenclature. The dotted line indicates the 15–48 tertiary interaction that is normally G-C or A-U. (A) Arrows point to base substitutions in the tRNA that either result in opal suppression activity (anticodon substitutions) or inactivate opal suppressor tRNA<sup>gly</sup> (C65 substitution). (B) Arrows point to the mutations that reactivate C65tRNA<sup>gly</sup>. Of the two Northern blot probes, one was complementary to positions 20–38 (endpoints marked by x), and the other was complementary to positions 29–47 (endpoints marked by o).
attempts, opal suppressor C65tRNAGly yielded too little DHFR to allow sequence analysis.

Activity of the Mutants in the Ribosome. I determined the suppression efficiency of the mutant tRNAs to assess their translational attributes. Suppression efficiency is the probability that the suppressor tRNA inserts its amino acid at UGA, thus allowing translating mRNA to continue producing nascent protein rather than terminating at the nonsense triplet. Each C65tRNA double mutant exhibited increased suppression efficiency relative to C65tRNA, consistent with the blue phenotype of the double mutants on the X-Gal plates. Cochella and Green (14) reported that a base substitution mutant in E. coli tRNATrp accelerates a specific step in the ribosome, which is in keeping with the increased translation activity of the mutants described here. The presence of residue A39 in one mutant may allow conversion of A37 to ms2i6A37 in the tRNA, which can improve the binding of the tRNA to the ribosome (15). I do not expect other mutants to alter tRNA modification.

Charging Levels in the Mutants. I determined the steady-state levels of glycyl-tRNA Gly in Northern blots of acid gels (Fig. 4 and Table 1). Acidic conditions stabilize the labile aminoacyl bond (16). Virtually all double and single tRNA mutants elevated charging relative to C65tRNA. The acid gels reveal differences in mobility and RNA levels among the mutant tRNAs. Neither difference affects the charging measurements because the percentage aminoacyl-tRNA is independent of the cellular concentration of the tRNA over a 75-fold range (17). Reanalysis of the same RNA preparations under fully denaturing and deacylating conditions revealed all tRNA bands migrated similarly to one position (Fig. 3). Therefore, it is likely that multiple conformational and dynamic properties of the mutant tRNAs produce the distinctive migration profiles in acid gels. Table 1 reports the relative tRNA yields in the denaturing gel calculated by dividing the Northern blot tRNA signal by that of 5S RNA in the same sample. For the double mutants, the yield varied 4.7-fold, whereas that of the single mutants varied 5.5-fold.

Relationship of Aminoacylation and Translational Efficiency. I investigated the correlation between glycyl-tRNA and suppression efficiency. A positive correlation would be expected if the tRNAs that efficiently suppress UGA also efficiently charge the tRNAs. Table 1 reports the values derived from quantitation of the RNAs in the gel. The apparent low yield of 5S RNA in the pGFIB sample in B is in part an artifact because to orient and identify lane 1 of the gel the corner of the hybridization membrane was cut off, which partially excised this 5S material. Further, the original data of the pGFIB sample show low-level signals migrating in the positions of acyl and deacyl cellular tRNA Gly.

### Table 1. Functional properties of the mutants

<table>
<thead>
<tr>
<th>Opal suppressor tRNA Gly gene</th>
<th>Gly in DHFR, %</th>
<th>Suppression efficiency, %</th>
<th>Glycyl-tRNA observed, %</th>
<th>Relative tRNA (tRNA Gly /H20862 5SRNA)</th>
<th>Relative glycyl-tRNA units</th>
</tr>
</thead>
<tbody>
<tr>
<td>C65</td>
<td>—</td>
<td>0.9</td>
<td>14.8**</td>
<td>0.54</td>
<td>27.4</td>
</tr>
<tr>
<td>U65*</td>
<td>98</td>
<td>4.5</td>
<td>33.3</td>
<td>0.52</td>
<td>64.0</td>
</tr>
<tr>
<td>C65U13</td>
<td>90</td>
<td>4.6</td>
<td>31.4</td>
<td>0.61</td>
<td>51.5</td>
</tr>
<tr>
<td>U65U13</td>
<td>98</td>
<td>9.2</td>
<td>33.8</td>
<td>0.58</td>
<td>58.3</td>
</tr>
<tr>
<td>C65U27</td>
<td>95</td>
<td>2.2</td>
<td>35.4</td>
<td>0.69</td>
<td>51.3</td>
</tr>
<tr>
<td>U65U27</td>
<td>96</td>
<td>5.4</td>
<td>41.8</td>
<td>0.24</td>
<td>174.2</td>
</tr>
<tr>
<td>C65A28</td>
<td>96</td>
<td>2.9</td>
<td>13.2**</td>
<td>1.44</td>
<td>9.2</td>
</tr>
<tr>
<td>U65A28</td>
<td>96</td>
<td>5.3</td>
<td>48.3</td>
<td>0.62</td>
<td>77.9</td>
</tr>
<tr>
<td>C65A39</td>
<td>94</td>
<td>4.9</td>
<td>34.5</td>
<td>1.0</td>
<td>34.5</td>
</tr>
<tr>
<td>U65A39</td>
<td>95</td>
<td>8.8</td>
<td>43.7</td>
<td>1.1</td>
<td>39.7</td>
</tr>
<tr>
<td>C65U48</td>
<td>96</td>
<td>2.1</td>
<td>43.0</td>
<td>0.81</td>
<td>53.1</td>
</tr>
<tr>
<td>U65U48</td>
<td>93</td>
<td>6.4</td>
<td>40.4</td>
<td>0.86</td>
<td>47.0</td>
</tr>
<tr>
<td>C65U75</td>
<td>92</td>
<td>2.8</td>
<td>19.5**</td>
<td>0.33</td>
<td>59.1</td>
</tr>
<tr>
<td>U65U75</td>
<td>96</td>
<td>5.0</td>
<td>41.8</td>
<td>0.22</td>
<td>190.0</td>
</tr>
</tbody>
</table>

**Materials and Methods describe details of the assays.

*I observed no amino acid other than Gly at residue 3 in DHFR; the background level was 5%.

1 Determined by PAGE at pH 5 in 7 M urea at 4°C followed by Northern blotted.

2 The fraction tRNA/5SRNA obtained after running the same samples on a fully denaturing gel and Northern probing. In the latter analyses, the C65 series and U65 series were run and probed separately.

3 The ratio of glycyl-tRNA observed/relative tRNA.

4 Insufficient protein to analyze.

5 Value similar to that with cells without an opal suppressor tRNA gene.

6 **Polydisperse gel species observed migrating more rapidly than deacyl-tRNA after acid urea PAGE.

7 Gene for opal suppressor derived from tRNA Gly (GCC anticodon) reported previously and included here for comparison.
efficiency by calculating the relative glycyl-tRNA (Table 1), i.e., dividing column glycyl-tRNA by relative tRNA. Fig. 5 shows a scatter plot of glycyl-tRNA units vs. β-galactosidase units. The plot of neither the double (Fig. 5A) nor single (Fig. 5B) mutant tRNAs indicates a straightforward correlation between the variables. However, the single mutants form two groups. Group one contains five members on the left, and group two contains U27 and U75 on the right. The groups separate from one another on the x axis, indicating mutually exclusive aminoacylation properties between members of the respective groups. Furthermore, whereas U27 and U75 occupy a similar x,y coordinate, the members of group one vary substantially on the y axis, suggestive of different efficiencies of ribosomal function among this group. The disparate scatter of data points in Fig. 5 indicates that mutations are idiosyncratic with respect to the mechanistic step affected.

U27 and U75 are the most efficient glycine acceptors on a per-molecule basis, but they are present in relatively low abundance and consequently are not the highest-efficiency suppressors (Fig. 5). With respect to translation and the peptidyl transferase center, U75 could potentially form U-G wobble pairs with the receptor G residues in the A (G2553) and P (G2251) sites of 23S rRNA that normally form a W-C pairs with C75 (18, 19). There could be two opposing effects of the U75 interactions. The A site interaction could be disfavored relative to that with C75 because one less hydrogen bond forms with the wobble pair relative to the base pair. Subsequently, a less stable interaction could facilitate translocation from A to the P site and the release from the P site. Overall, weaker binding of U75 to the rRNA may be one reason this tRNA is not among the high-efficiency suppressors.

The reduced cellular level of U75tRNA may reflect enhanced RNase T degradation specific for uncharged tRNA with a CUA-3′ end, which would produce tRNA(C74-3′) (20). It is conceivable, although not experimentally demonstrated in vivo, that tRNA(C74) is susceptible to further degradation to smaller products (21). Curiously, tRNA(C74-3′) is not efficiently repaired to tRNA(CCA-3′) by tRNA nucleotidyltransferase (22). Perhaps the ribonucleases are compartmentalized in an unknown E. coli degradation complex. Alternatively, although there is no evidence for it, a transcriptional defect could reduce the amount of U75tRNA.

As an aminoacylation reference, I determined the level of glycyl-tRNA(Gly) for WT tRNA(Gly) (GCC anticodon). This measurement yielded 75.8% in E. coli XAC/U4 cells carrying plasmid pGFIIB without an inserted tRNA gene and 73.8% in these cells expressing tRNA(Gly) (GCC anticodon) from pGFIIB. Thus, changing the anticodon of tRNA(Gly) from GCC to UCA (Table 1, glycyl-tRNA observed for U65) reduces glycylation 2.2-fold. However, despite lacking anticodon determinant C36 (10), the A28 single mutant of opal suppressor tRNA(Gly) produces 48.3% glycyl-tRNA (Table 1 glycyl-tRNA observed), which is just 1.5-fold less than that of WT tRNA(Gly). The latter result points to the superiority of the compensating mutations.

Bioinformatic Analysis of tRNA(Gly). Fig. 6 shows the number and type of mispairs at various base pair positions in the subset of all tRNA(Gly) sequences (n = 91) present in the 1,550 data set of 20 accepting types (23). Mispairs at positions 49–65, 10–25, and 13–22 predominate; moreover, they are enriched an average of 2.6-fold in tRNA(Gly) relative to mispairs in the entire 1,550 data set (cf. Fig. 1). An additional distinguishing feature of tRNA(Gly) is that >50% of its positions contain ≤1 mispair. By contrast, just 10% of the positions in the larger data set contain ≤1 mispair after normalizing for different numbers of sequences in the individual acceptor types. In summary, the profile of mispairs in tRNA(Gly) differs markedly from that of the larger data set. Genomic tRNA(Gly) sequences that reiterate individual mutations contain multiple glycine anticodons. The overall mispair profile of genomic tRNA(Gly) is a distinctive signature of the molecule.
with an evident functional correlate uncovered by the genetic analysis.

**Relationship of Mutational Positions and Statistically Hot Positions.**
Remarkably, the positions of compensating mutations generating mispairs in *E. coli* tRNA<sub>Gly</sub> match the genomic pattern of mispairs in the 91 genomic tRNA<sup>Gly</sup> sequences (compare Figs. 2B and 6). Fisher’s exact test of independence measures the significance of this type of overlap (24). If the two processes (i.e., the production of compensating mutations and nature’s genomic mispairs) are unrelated, the chance of seeing as much overlap is \( P = 0.025 \). I note two further points about the genomic mispairs. First, with one exception (31–39), the mutational positions have many repeats in genomic tRNA<sup>Gly</sup> sequences. Second, position 10–25 has many mispairs in genomic tRNA<sup>Gly</sup> but no compensating mutation. Because the mutagenesis was not statistically saturating, I constructed two mutants of opal suppressor tRNA<sup>Gly</sup> containing this mispair. Coupling the G-U with C65 gave 19.5% aminoacyl-tRNA<sup>Gly</sup> observed by acid gel analysis. The other mutant that combined the query G-U with G-U at position 49–65 yielded 31.3% aminoacyl-tRNA<sup>Gly</sup>; I measured neither the glycine specificity nor the suppression efficiency of these mutants so their functional relationship to the group of mutants has not been determined. Finally, I cannot exclude the possibility that all of the genomic mispairs, regardless of frequency, are members of the group or that they are members of another related group.

It is interesting to note that because the six mispairs of U65tRNA<sup>Gly</sup> are more active than the normal suppressor tRNA, then tRNA<sup>Gly</sup> genomic sequences containing G-U at 49–65 should reiterate the group mispair pattern; analysis confirms this corollary (results not shown). Second, G-C/C-G and/or A-U/U-A at 49–65 should exhibit a propensity for the same associated mispairs. Such pairs in tRNA<sup>Gly</sup> are prominent with G-C/C-G, but not with A-U/U-A (results not shown). The positional specificity of mispairs in G-U/U-G and G-C/C-G tRNAs relative to the mutational profile is remarkable and greatly reinforces the general correctness of my hypothesis.

**Conclusion**
Genetic and bioinformatic analyses of tRNA<sup>Gly</sup> have identified a shared set of mispair sites that enhance the molecules’ aminoacylation and translation functions. By analogy with prior biophysical work in other RNA system (25–28), the mutant tRNAs probably are more active because their mispairs exhibit greater conformational heterogeneity and thermodynamic instability, whereas inactive W-C pairs and their helices are structurally rigid. Thus, the deformability of the mutant tRNAs allows their adaptability to the dynamics of synthetases and translation mechanisms at lower free energy costs relative to rigid W-C helices. The particular mechanistic step enhanced by individual mispairs is probably idiosyncratic. Further, I suspect that the two cognate pairs of the C65 double mutants, the tRNA and synthetase, and the charged tRNA and ribosome, each behave as overall shape or flexibility units that sum their effects (Fig. 5A). Conversely, the single mutants appear to behave as independent domains that do not sum their effects (Fig. 5B). Clearly, further analysis of these and related mutants would be valuable.

**Materials and Methods**
**Construction of Opal Suppressor tRNA and Isolation of Strains with Compensating Mutations.** I derived mutants of opal suppressor tRNA genes from corresponding mutant amber suppressor genes once I learned that a number of the latter mutants had mixed amino acid specificities. All mutant opal suppressors were glycine-specific except C65.

Measurements of parameters in the opal suppressor system were performed on 37°C cultures that had been inoculated with freshly transformed clones (9, 29). The DHFR gene contained a UGA triplet in place of codon 3, and the percent amino acid yield at residue 3 in the protein sequence is reported. Amino acid values < 5% are not reproducible and are not reported but are retained in the calculation, so none of the yields sums to 100%.

**Suppression efficiency values were determined in *E. coli* XAC/U4 cells that harbor the lacI-Z fusion system carrying UGA nonsense allele U4. Values are the average of triplicate measurements and are reported as the percentage of mutant enzyme activity relative to that of I-Z40 fusion, which carries the WT amino acid codon in place of the U4 triplet. The values have not been corrected for the value of cells without an opal suppressor tRNA, which was 1.0%.

**Mutagenesis.** I mixed plasmid DNA in 300 μl of 0.1 M phosphate buffer (pH 6) with 200 μl of 1 M hydroxylamine solution and incubated at 50°C for 4 h. After overnight dialysis at room temperature against 10 mM Tris-HCl (pH 8)/40 mM sodium acetate, the DNA was precipitated overnight at −20°C in 2 vol of absolute ethanol, collected by centrifugation, and resuspended in 20 μl of Tris-HCl (pH 8)/1 mM EDTA. Cells were transformed with 2 μl of the treated DNA solution, and serial dilutions were spread on X-Gal agar plates.

**Polyacrylamide Gel Analysis of tRNAs.** I determined the relative amounts of opal suppressor tRNA<sub>Gly</sub> by RNA fractionation on a completely denaturing polyacrylamide gel followed by Northern blot analysis. RNA samples (~0.1 OD<sub>260</sub> per lane) in formamide loading buffer were applied to a gel that contained 8% polyacrylamide, 7 M urea, 89 mM Tris-borate (pH 8.3), and 2 mM EDTA and were run at 55–60°C at 1,500 V. The locations of two tRNA probes are shown in Fig. 2B. Each mutant tRNA was capable of binding one probe or the other with perfect comple-

---

*Fig. 6.* Histogram showing the frequency of G-U and other mispairs at various base pair positions in 91 tRNA<sup>Gly</sup> sequences. The arrows mark base pair sites where compensating mutations occur.
mentarity (i.e., no mismatches), and, because both probes overlap the anticodon loop–stem region, a given tRNA could not bind two probes simultaneously. The altered anticodon of the tRNAs ensure the probes hybridize the suppressor tRNA in great preference to WT chromosomal tRNA. An internal standard, 5S RNA, was probed simultaneously with an oligonucleotide complementary to positions 34–53. All probes were 32P-labeled at their 5’ ends. Table 1 reports the values derived from quantitation of the RNAs using a Typhoon phosphorimagere and IMAGEQUANT software (Amersham Biosciences).

Aminoacyl-tRNA and deacyl tRNA were isolated with bulk low-molecular-weight RNA from cells under acidic conditions (sodium acetate, pH 5.2) and ~0.1 OD260 per lane were separated on partially denaturing 6.5% polyacrylamide gel containing 8 M urea (pH 5.2) run at 4°C and 500 V. tRNAs were detected by a Northern blot analysis using the two described tRNA probes. Calculation of the steady-state level of glycyl-tRNA as follows. I based the calculation on the 21 sites (21 helical base pair sites plus 15–48) that differ from the target site (49–65). I note that 5 of the 21 sites exhibited mispairing in the mutagenesis experiment. Further, 7 of the 21 sites exhibited frequent mispairing throughout evolution (i.e., counts in Fig. 6 where the value is >5). It is striking that four base pair sites are in common between the two sets of sites (13–22, 15–48, 27–43, and 28–42). I used Fisher’s exact test of independence (24) to measure the significance of this overlap.

I thank Glenn Björk, Michael Newton, Gabriele Varani, Wolfgang Wintermeyer, and Mike Yarus for discussions and Hynsic Choi, Kay Gabriel, Jay Schneider, and Po-Sheng Tsai for technical assistance. This work was supported by National Institute of General Medical Sciences Grant GM42123, the College of Agricultural and Life Sciences at University of Wisconsin-Madison, and the State of Wisconsin.

Probability Calculation. I calculated the significance of the agreement between compensating mispairs and nature’s genome mispairs in tRNA\textsuperscript{Gly} as follows. I based the calculation on the 21 sites (21 helical base pair sites plus 15–48) that differ from the target site (49–65). I note that 5 of the 21 sites exhibited mispairing in the mutagenesis experiment. Further, 7 of the 21 sites exhibited frequent mispairing throughout evolution (i.e., counts in Fig. 6 where the value is >5). It is striking that four base pair sites are in common between the two sets of sites (13–22, 15–48, 27–43, and 28–42). I used Fisher’s exact test of independence (24) to measure the significance of this overlap.