Optimizing the spiking fraction

The three cases GG, CC, and GC are particularly challenging to distinguish directly, in which case 'spiking', or adding a certain fraction of a known homozygote, can provide additional information in terms of the heteroduplexes which may be formed in the resulting mixture. Most simply, if the unknown sample is identical to the known spike, there will be no heteroduplexes formed and the melting curve does not change. If the unknown sample is the opposite homozygote, then the spike will cause some heteroduplexes will be formed. If the unknown sample is heterozygous, the presence of the spike will reduce the heteroduplex proportion from one half of all duplexes (one fourth for each of two types) to some smaller fraction.

We wish to optimize the amount of spike we should add, to best separate the heteroduplex proportions resulting from mixing it with each of these three unknowns. If we use x to denote the fraction of spike in the mixture then y = 1 - x is the fraction of unknown sample. For example if we add $\frac{1}{3}$ as much spike as unknown, then $x = \frac{1}{4}$ and $y = \frac{3}{4}$. The total heteroduplex fraction for each of these three relationships of the unknown to the spike is given by

$$h_1(x) = 0$$

$$h_2(x) = 2xy = 2x(1-x) = -2x^2 + 2x$$

$$h_3(x) = 2x'y' = 2(x + \frac{1-x}{2})(\frac{1-x}{2} = 2(\frac{1+x}{2})(\frac{1-x}{2} = \frac{1-x^2}{2} = -\frac{1}{2}x^2 + \frac{1}{2}$$

The formula for h_1 states that if the spike and unknown represent the same homozygote, not heteroduplexes are formed. The formula for h_2 states that two kinds of heteroduplexes are formed, each in proportion to the fraction of the complementary strands present. The proportion of each strand of spike is x and each strand of unknown is y. Spike homoduplexes are formed in proportion $xx = x^2$, unknown homoduplexes are formed in proportion $yy = y^2$, and two types of spike-unknown heteroduplexes are formed in proportion xy + yx. The same reasoning applies to h_3 with the understanding that half of each of the strands comprising the unknown are the same as those of the spike and half are opposite. Therefore we divide the unknown portion y = 1 - x into two halves, one of which we combine with the spike $x' = x + \frac{1-x}{2}$, and the remainder, $y' = \frac{1-x}{2}$, and compute the proportion of heteroduplexes formed, x'y' + y'x', as we did in the previous case.

Now we need only look to find where the separation between these three graphs is maximized. Observe that h_3 has $\frac{1}{2}$ heteroduplexes while h_1 and h_2 have zero heteroduplexes when no spike is present (x = 0) and all three have zero heteroduplexes when there is nothing but spike (x = 1).

The graphs of h_2 and h_3 intersect at $x = \frac{1}{3}$, where $h_2(\frac{1}{3}) = h_3(\frac{1}{3}) = \frac{4}{9}$. So the best separation will be for some value of $x < \frac{1}{3}$ where h_2 is equidistant from $h_1 = 0$ and h_3 . This observation leads to the equation $h_3 - h_2 = h_2 - h_1$ or

$$-\frac{1}{2}x^2 + \frac{1}{2} - (-2x^2 + 2x) = -2x^2 + 2x$$

which is equivalent to the standard form

$$\frac{7}{2}x^2 - 4x + \frac{1}{2} = 0$$

The quadratic formula gives two solutions, x = 1 (where we have observed $h_1 = h_2 = h_3 = 0$) and the optimal spiking fraction, $x = \frac{1}{7}$. Our conclusion is that adding 1/6 as much spike as unknown, resulting in proportions $x = \frac{1}{7}$ and $y = \frac{6}{7}$, provides the optimal separation in the proportion of heteroduplexes in the resulting mixtures. In particular, $h_1(\frac{1}{7}) = 0$, $h_2(\frac{1}{7}) = 12/49$, and $h_3(\frac{1}{7}) = \frac{24}{49}$.

The analysis does not depend upon the base in the mutant homozygote being complementary to the base in the wild type, i.e., the same spike fraction would also optimal for distinguishing among, for example, GG and AA homozygotes and AG heterozygotes.