Supplementary Figure Legends

Figure S1. Graphical presentation of (A) a conventional histogram based on bins (buckets), and (B) an analogous histogram derived from digital pH curve points.

A) The number of individual observations associated with each bin x_k (from k = 1 to k = m) corresponds to weight W_k (in this example: W_1 = 2; W_2 = 4; W_3 = 5; ....; W_m = 1). Each rectangle represents an individual observation (= individual contribution to the distribution function). The total number of observations n is the sum of all weights: $n = \sum_{k=1}^{m} W_k$.

B) The intensities I_k (arbitrary unit) of digital pH curve points pH_k (from k = 1 to k = m) correspond to weights W_k (in this example: $I_1 : I_2 : I_3 .... : I_m = W_1 : W_2 : W_3 .... : W_m$ = 2 : 4 : 5 .... : 1, by analogy to (A)). $n = \sum_{k=1}^{m} W_k$.

Figure S2. Conversion of a ^{31}P NMR spectrum of 3-APP into a pH_e distribution curve. Positions and intensities (weights) of digital points are indicated by vertical lines in (A) to (C), and by symbols in (D). This analysis is based on the spectrum shown in Fig. 1 B; however, for better visibility of the vertical lines, ppm and pH ranges were slightly reduced in (A) to (C).

A) 3-APP spectral region before ppm-to-pH conversion. Note the evenly spaced vertical lines representing digital spectrum points. The underlying ^{31}P NMR spectrum was obtained from a CCL39 mouse tumor. The height of each digital point k is denoted by I_k (intensity).

B) Uncorrected pH_e distribution obtained by ppm-to-pH conversion of (A), resulting in unequal intervals between digital points, due to the nonlinear relationship between the ppm and pH scales.

C) Corrected pH_e distribution, obtained from (B) by correcting the intensities of digital points for effects introduced by the nonlinear relationship between the ppm and pH scales. The

resulting intensities (weights) are denoted by I_k^{corr} (= W_k). The envelope of this pH distribution represents the final pH curve (pH profile).

D) Overlay of the digital points representing the pH curve (C) (triangles), and the same curve points after intensity (= weight) rescaling as needed for accurate statistical parameters. Circles, intensities rescaled by interval weighting; diagonal crosses, intensities rescaled by analytical weighting. The analytically rescaled pH distribution is equivalent to the distribution presented in (B).

E) Effects of converting Gaussian spectral lines (top) to pH curves (bottom). Two ideal 3-APP ^{31}P NMR resonances were numerically simulated to result in pH modes characteristic of tumors: pH_{e1} = 7.2 and pH_{e2} = 6.5, *i.e.* $pH_{e1} > pK_{a2}$, and $pH_{e2} < pK_{a2}$. Symmetric and rather broad Gaussian spectral lines yielded relatively narrow and asymmetric pH distributions (enhanced tails extending away from pK_{a2} of 3-APP; empty arrows indicate unilaterally broadened curve bases). Note that the tails of the pH profiles (bottom) are considerably lighter than the tails of the underlying spectral lines (top), in agreement with measured data (Fig. 4 C *vs.* A). This suggests that the extended tails of the spectral lines do not contribute significantly to define the actual pH distribution curves. Gaussian lineshapes were chosen because our experimental spectral lines have a strong Gaussian character due to the routinely used Lorentzian-Gaussian lineshape transformation (for statistical results see Table S2).

F) Hypothetical 3-APP ^{31}P NMR signals (top) calculated from simulated, perfectly Gaussian pH distributions (bottom) for pH modes 6.5 and 7.2. Symmetric and rather narrow Gaussian pH distributions yield relatively broad and asymmetric spectral lines. For statistical results see Table S2 (unimodal distribution) and Table S3 (bimodal distribution).

For optimal visualization of asymmetries in transformed curves, linewidths were set to values exceeding those commonly found in *in vivo* or phantom spectra. Kurtosis was generally decreased (both as a signed and as an absolute number) in pH curves derived from

measured *vs.* simulated (Gaussian) 3-APP ^{31}P NMR spectral lines. Thus, measured pH curves are both flatter and more Gaussian-like than these simulated pH curves, although measured spectra were processed with Lorentzian-Gaussian lineshape transformation. This indicates that *in vivo* spectral lineshapes of 3-APP and P_i were not predominantly Gaussian, but were indeed determined by pH heterogeneity that had some Gaussian character. However, positive or negative skewness was still significant in many *in vivo* pH curves, indicating asymmetric pH distributions.

Figure S3. Scaling corrections.

(A) to (D) Conversion of a bimodal pH_e distribution obtained from a ^{31}P NMR spectrum of 3-APP to a pH curve usable for parameter fitting to two overlapping analytic curves. Positions and intensities (weights) of digital points are indicated by vertical lines. This analysis is based on the spectrum shown in Figs. 1 B and S2 A - D.

A) pH_e distribution with unequal intervals between digital points.

B) Distribution diagram derived from (A) with equalized intervals between digital points.

C) Distribution diagram obtained from (B) through nonlinear rescaling of the weights of digital points for effects introduced by equalized intervals. The shape of this curve is identical to that of the underlying ^{31}P NMR spectral region.

D) Same distribution diagram as (C), displayed with an increased spectral range for best curve fitting.

(E) to (J) Effects of n on nominal standard deviation (s), skewness (G1) and kurtosis (G2) of a modeled normal (Gaussian) pH distribution for m = 11 distribution curve points (E to G), and of a bimodal pH distribution based on a phantom 3-APP ^{31}P NMR spectrum (Fig. 4 I) for m = 38 (H to J). In our statistical analysis of pH heterogeneity, the number m of abscissa values (analogous to bins) is determined by the number of digital points of the spectral region

used, and n is proportional to m and the weights W_k. Each W_k is, in turn, proportional to the intensity of spectrum point k and arbitrarily scalable. These diagrams show that with increased linear upscaling of W_k, n approaches values that are sufficiently high (n>>m) to render skewness and kurtosis independent of the scaling factor chosen. Consequently, s, G1 and G2 asymptotically approach constant values. Our algorithms and the corresponding algorithms provided by EXCEL yielded identical results for s, G1 and G2.

The sensitivity of skewness and kurtosis to the number m of a typical pH curve was tested for the pH_e distribution presented in panel G. Halving m by omitting every other point had negligible effects (≤ 0.03 pH units) on $\overline{pH_e}$, $\widetilde{pH_e}$, pH_{e1}, pH_{e2}. and pH_{e2}/pH_{e1} area ratios, but minor changes were observed for G1 (from 0.34 to 0.40), G2 (from -0.38 to -0.35), and H (from 6.14 to 5.13). As a consequence, curve shape parameters can be compared to good effect between pH profiles with differing m, although for best reproducibility constant m values are preferable.

Figure S4. Simulation of bimodal pH distributions by addition of computer-generated Gaussian curves. The contribution of the pH_{e2} region to the overall distribution was systematically varied by changing the relative intensities of these curves, using the intensity ratios given in Table S3. The vertical dashed line in each of the curves (A) to (E) indicates the integration limit, *i.e.* the minimum near the center of the pH distribution. For (F), the pH_{e2} contribution is virtually undetectable.

Diagram (G) shows the integration error for the area under the pH_{e2} region (in percent of the theoretical value) as a function of decreasing pH_{e2} contribution, resulting in increasingly underestimated relative pH_{e2} areas (G). This error considerably exceeded 10% only when the pH_{e2} contribution to the pH distribution decreased below 20% of the pH_{e1} contribution (*i.e.* when the pH_{e2} contribution approached 10% of the overall pH_e area). This threshold presents the lower limit for separate quantification of distinct pH regions (see RESULTS). The origin of

this error is the shift of the integration limit (curve minimum between the pH_{e1} and pH_{e2} regions) toward pH_{e2} as the relative contribution of the pH_{e2} region to the pH profile decreases (Table S3 panels A to E).

Figure S5. Gaussian curve fitting to 3-APP ^{31}P NMR spectral regions of phantoms and mouse tumors for varying numbers and intensities of pH_e modes.

(A) to (D) Bimodal spectral regions of phantoms. Low-pH (left mode) 3-APP concentrations decrease from left to right, corresponding to the pH distributions in Figure 4 I - L. These distributions were fitted very well, as demonstrated by the perfect superimposition of the measured distribution function (symbols) and the numerical fit (solid lines).

E) Trimodal spectral region of a 3-APP phantom, corresponding to the pH distribution in Figure 4 H.

(F) to (H) Bi- and multimodal spectral regions from 3-APP-injected mouse tumors, corresponding to the pH distributions shown in Figs. 3 D, 1 A, and Fig. 3 H, respectively. Panel F represents the pH_e profile of a tumor with a very clean bimodal distribution which was fitted very well by a combination of two Gaussian-Lorentzian curves. By contrast, panel G shows an apparent bimodal distribution that was best fitted by a combination of three Gaussian-Lorentzian curves. Panel H illustrates that curve fitting of irregularly shaped pH distributions may not be meaningful as the number of peaks needed for a good fit would have to be high and rather arbitrary, and the underlying peak shapes would be ill-defined.

Black symbols represent measured digital spectrum points, solid lines represent fitted curves. Arrows indicate the location of the curve maxima representing pH modes. These maxima were chosen for optimal curve fitting.

Figure S6. Flowchart presenting the step-by-step procedure for calculating multiple pH heterogeneity parameters from ^{31}P NMR spectra.

The left column shows the procedure used for ^{31}P NMR spectra directly transferred to our EXCEL spreadsheet. The right column displays the initial steps required if the spectral region of interest is to be fitted by an external program (*e.g.*, spectrometer software) before processing the resulting curves with our EXCEL spreadsheet. In the latter case, the external program in question yields fitted parameters that are utilized by our EXCEL spreadsheet to construct curves for further processing. Alternatively, external fit programs that produce a fitted curve of the entire spectral region of interest can be employed. In such cases, the steps contained within braces are to be omitted.