$$I\left(\Delta N_i; R_J\right) = I\left(N_i; R_J\right) - I\left(n - N_i; R_J\right) \qquad \text{Equation S1}$$

$$I\left(N_i; R_J\right) = \log\left(\frac{f_{N_i,R_j}/f_{R_j}}{f_{N_i}/R}\right) \qquad \text{Equation S2}$$

$$I\left(n - N_i; R_J\right) = \log\left(\frac{f_{n-N_i,R_j}/f_{R_j}}{f_{n-N_i}/R}\right) \qquad \text{Equation S3}$$

$$I\left(\Delta N_i; R_J\right) = \log\left(\frac{f_{N_i,R_j}}{f_{n-Ni,Rj}}\right) + \log\left(\frac{f_{n-N_i}}{f_{N_i}}\right) \qquad \text{Equation S4}$$

S1-S4 Equations: Basis for the development of the RRMScorer equation. Our method is based on the information difference between the occurrence of two events, in our case the information of how often a specific nucleotide interacts with a specific residue, $I\left(N_i; R_J\right)$, and the information when that same nucleotide interacts with any other residue, $I\left(n - N_i; R_J\right)$ (equation S1). The individual terms are developed in equations S2 and S3 where: f_{N_i,R_j} is the number of occurrences for a specific contact between nucleotide *i* and residue *j*; f_{R_j} the number of times residue *j* is in that position; f_{N_i} the number of times nucleotide *i* is in that position; f_{n-N_i,R_j} the number of times nucleotide *i* interacts with any residue but residue *j*; f_{n-N_i} the number nucleotides other than nucleotide *i* in that position; R the total number of residues in the dataset. As it's a difference between the two logarithms, the common terms that account for the number of specific residues in position *j* (f_{R_j}) and total number of residues in the dataset (R) disappear from the equation. After the simplification we obtain equation S4 (Same as equation 2 in the main text).