



Figure 7. Acceptance probabilities for samples on unseen test peptides with the Timewarp MCMC algorithm. The red line is at 0.01%, below that efficient sampling becomes difficult.

Hence  $p$  is translation equivariant even if  $q$  is not.

### A.3. Chirality

In addition to the symmetries described in Section 4.1 the potential energy  $U(x^p)$  of a molecular configuration is also invariant under mirroring. However, in the presence of *chirality* centers, a mirrored configuration is non-superposable to its original image (Kelvin, 1894). An example of a chirality center in an amino acid is a Carbon atom connected to four different groups, e.g. a  $C_\alpha$  atom. In nature most amino acids come in one form, namely L-amino acids. Hence, all our datasets consist of peptides containing only L-amino acids. In rare cases, as the model proposes large steps, one step might change one L-amino acid of a peptide to a D-amino acid in a way that the resulting configuration has a low energy and the step would be accepted. We prevent this by checking all chirality centers for changes at each step and reject samples where such a change occurs. This does not add any significant computational overhead.

## B. Additional results

In this section we show additional results like the conditional distribution as well as more peptide examples for experiments discussed in Section 6.

### B.1. 2AA additional results

More examples from the 2AA test set are presented in Figures 8 and 9. We achieve the worst speed-up for the dipeptide GP (Figure 8 last row) as it does not show any slow transitions. The distribution of the acceptance probabilities for the Timewarp MCMC algorithm is shown in Figure 7.

### B.2. 4AA additional results

More examples from the 4AA test set are presented in Figures 10 and 12. The distribution of the acceptance probabilities for the Timewarp MCMC algorithm is shown in Figure 7.

### B.3. Conditional distributions

The model was trained to generate samples from the conditional Boltzmann distribution  $\mu(x(t + \tau)|x(t))$ . Here we show some examples of the conditional distribution generated by the Timewarp model compared to the conditional distribution induced by MD. While we can generate 5,000 samples from the conditional distribution of the model in parallel, we require 5,000 distinct MD trajectories of simulation length  $\tau$  to sample the conditional distribution with MD. Hence, generating samples from the conditional distribution is several orders of magnitude faster with the model. In Figures 11 to 13 we show example conditional distributions for alanine dipeptide and peptides from the 2AA and 4AA datasets. For all peptides the model learns a conditional distribution that is close to the conditional MD distribution. Moreover, the relative weights in the TICA projections as well as the bondlength distributions match very well. Only the energies of the model samples are