

Enhanced Sampling of Configuration and Path Space in a Generalized Ensemble by Shooting Point Exchange

Sebastian Falkner, Alessandro Coretti, and Christoph Dellago*
University of Vienna, Faculty of Physics, 1090 Vienna, Austria.
 (Dated: March 23, 2023)

The computer simulation of many molecular processes is complicated by long time scales caused by rare transitions between long-lived states. Here, we propose a new approach to simulate such rare events, which combines transition path sampling with enhanced exploration of configuration space. The method relies on exchange moves between configuration and trajectory space, carried out based on a generalized ensemble. This scheme substantially enhances the efficiency of the transition path sampling simulations, particularly for systems with multiple transition channels, and yields information on thermodynamics, kinetics and reaction coordinates of molecular processes without distorting their dynamics. The method is illustrated using the isomerization of proline in the KPTP tetrapeptide.

Overcoming high energy barriers to explore configuration and trajectory space in simulations of rare events is at the core of the sampling problem. Numerous enhanced sampling techniques have been developed over the years to better understand the thermodynamics and kinetics of rare events such as nucleation, chemical reactions and biomolecular reorganization [1–5]. However, when deciding for the most suitable method, a conflict of interest frequently arises. Enhanced sampling methods such as metadynamics [6, 7] and umbrella sampling [8] efficiently focus computational resources on the regions of interest while still allowing for reweighting to gain information on the equilibrium distribution $p_{\text{eq}}(x)$. However, the dynamics of the system are distorted by the introduction of the bias potential. In contrast, path sampling methods such as Transition Path Sampling (TPS) and Transition Interface Sampling (TIS) allow to obtain true-dynamic trajectories between stable states [9, 10]. However, these schemes may suffer from correlations between subsequently sampled trajectories, in particular for systems with several reaction channels. To alleviate this problem, in recent years there has been great interest in applying enhanced sampling methods to path space [11–14]. Apart from that, points on the trajectory are not distributed according to $p_{\text{eq}}(x)$. Although in principle possible, it is unfeasible to reweight samples to the equilibrium distribution since this requires knowledge of the committer probability of each point [15].

In this letter, we propose a sampling scheme based on a parallel sampling of configuration and path space, e.g., using metadynamics and TPS. The two simulations are coupled by exchanging configurations between them following an acceptance criterion derived for a generalized ensemble. As a result, transition paths show less correlations due to fast relaxation in configuration space and barriers in configuration space are crossed more frequently due to exchanges with configurations on transition paths.

Exchange moves, e.g. as employed in replica exchange molecular dynamics [16, 17] or replica exchange TIS [18], are a powerful tool to enhance sampling and reduce correlations between samples. In such a move [19], configurations of two systems are exchanged according to an acceptance criterion to ensure a properly weighted ensemble in both systems in the limit of infinite sampling. We propose to perform exchange moves between configurations y from a given distribution $p_y(y)$ and a configuration on a path X from the distribution of transition paths $P_X^{\text{AB}}(X)$, which includes only paths that connect two given regions A and B. In this context, a generalized ensemble can be defined through the joint distribution:

$$P_z(z) = p_y(y) P_X^{\text{AB}}[X(\tau)] \quad (1)$$

where the state z is given by $z = \{y, X(\tau)\}$ with a configuration y and a path $X(\tau)$ of length τ . The probability density of reactive paths $P_X^{\text{AB}}[X(\tau)]$ is given by [10]:

$$P_X^{\text{AB}}[X(\tau)] = \frac{1}{Z_{\text{AB}}} H_{\text{AB}}(x_0, x_\tau) \prod_{i=1}^{\tau/\Delta t - 1} \tilde{h}(x_{i\Delta t}) \times P_X[X(\tau)] \quad (2)$$

where Δt is the timestep, $p(x_{i\Delta t} \rightarrow x_{(i+1)\Delta t})$ is the short-time transition probability from $x_{i\Delta t}$ to $x_{(i+1)\Delta t}$ and Z_{AB} is the partition function. The probability distribution of an unconstrained path $P_X[X(\tau)]$ is given by [10]:

$$P_X[X(\tau)] = p_{\text{eq}}(x_0) \prod_{i=0}^{\tau/\Delta t - 1} p(x_{i\Delta t} \rightarrow x_{(i+1)\Delta t}) \quad (3)$$

where $p_{\text{eq}}(x)$ is the equilibrium or stationary distribution for the underlying dynamics of X . The factor $H_{\text{AB}}(x_0, x_\tau)$ is one if the trajectory connects states A and B in any order and is zero otherwise. The function $\tilde{h}(x)$ is zero if x is in state A or B and unity otherwise, ensuring that the transition path has exactly one point in state A and one in B.

For an exchange between the two spaces in $P_z(z)$, Eq. (1), a new state $z' = \{y', X'(\tau')\}$ is generated based

* christoph.dellago@univie.ac.at

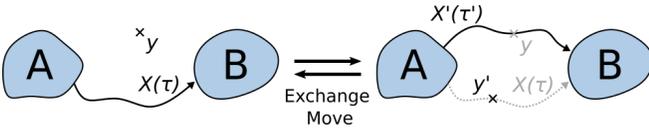


Figure 1. Schematic representation of the proposed exchange move for the generation of a new path $X'(\tau')$ and configuration y' given the path $X(\tau)$ and configuration y .

on the current state $z = \{y, X(\tau)\}$ (Figure 1). The new configuration y' is obtained by selecting a point on the current trajectory $X(\tau)$ with probability $p^{\text{sel}}[y'|X(\tau)]$. The generation probability for this move is given by:

$$p_y^{\text{gen}}[X(\tau) \rightarrow y'] = p^{\text{sel}}[y'|X(\tau)] \quad (4)$$

A new path X' is generated based on y by means of a shooting move where the equations of motion are integrated forward and backward in time until a stable state is reached. With $k\Delta t$ being the time of the shooting point on the new path, the generation probability is given by:

$$P_X^{\text{gen}}[y \rightarrow X'(\tau')] = \prod_{i=k}^{\tau'/\Delta t - 1} p(x'_{i\Delta t} \rightarrow x'_{(i+1)\Delta t}) \times \prod_{i=1}^k \bar{p}(x'_{i\Delta t} \rightarrow x'_{(i-1)\Delta t}). \quad (5)$$

Assuming that the transition probabilities fulfill microscopic reversibility, the above distribution can be rewritten as:

$$P_X^{\text{gen}}[y \rightarrow X'(\tau')] = \frac{1}{p_{\text{eq}}(y)} \times P_X[X'(\tau')] \quad (6)$$

Imposing detailed balance, the acceptance probability for the described exchange move must obey

$$\frac{P_z^{\text{acc}}(z \rightarrow z')}{P_z^{\text{acc}}(z' \rightarrow z)} = \frac{P_z(z') P_z^{\text{gen}}(z' \rightarrow z)}{P_z(z) P_z^{\text{gen}}(z \rightarrow z')} \quad (7)$$

which can be satisfied using the Metropolis rule;

$$P_z^{\text{acc}}(z \rightarrow z') = \min \left\{ 1, \frac{P_z(z') P_z^{\text{gen}}(z' \rightarrow z)}{P_z(z) P_z^{\text{gen}}(z \rightarrow z')} \right\} \quad (8)$$

Inserting expressions from Eq. (1), (4) and (6), the acceptance criterion for the exchange move is:

$$P_z^{\text{acc}}(z \rightarrow z') = H_{\text{AB}}(x'_0, x'_\tau) \prod_{i=1}^{\tau'/\Delta t - 1} \tilde{h}(x'_{i\Delta t}) \times \min \left\{ 1, \frac{p_y(y') p_{\text{eq}}(y) p^{\text{sel}}[y|X'(\tau')]}{p_y(y) p_{\text{eq}}(y') p^{\text{sel}}[y'|X(\tau)]} \right\} \quad (9)$$

The exchange scheme is most efficient if $p_y(y)$ has significant overlap with $p_{\text{eq}}(y)$ and shooting moves starting

from samples of y have a reasonable probability of generating a transition path. Both of these conditions can be matched well by setting the distribution $p_y(y)$ to the Boltzmann distribution with a bias potential introduced via metadynamics [6, 7]:

$$p_y(y) = Z^{-1} \exp\{-\beta[U(y) + U_{\text{bias}}(r(y))]\} \quad (10)$$

The acceptance of the exchange scheme can then be further improved by tuning the selection probability $p^{\text{sel}}[y'|X(\tau)]$, which represents the probability to choose a point y' on a given path $X(\tau)$. We can bias this selection in the spirit of Jung et al. [20] according to the current bias introduced by metadynamics:

$$p^{\text{sel}}[y'|X(\tau)] = \frac{\exp\{-\beta[U_{\text{bias}}(r(y'))]\}}{\sum_{i=0}^{\tau/\Delta t} \exp\{-\beta[U_{\text{bias}}(r(x_{i\Delta t}))]\}} \quad (11)$$

The acceptance probability for the exchange then becomes:

$$P_z^{\text{acc}}(z \rightarrow z') = H_{\text{AB}}(x'_0, x'_\tau) \prod_{i=1}^{\tau'/\Delta t - 1} \tilde{h}(x'_{i\Delta t}) \times \min \left\{ 1, \frac{\sum_{i=0}^{\tau/\Delta t} \exp\{-\beta[U_{\text{bias}}(r(x_{i\Delta t}))]\}}{\sum_{i=0}^{\tau'/\Delta t} \exp\{-\beta[U_{\text{bias}}(r(x'_{i\Delta t}))]\}} \right\} \quad (12)$$

The resulting criterion therefore represents the ratio of the times the old and new paths spend in regions with a high bias potential. This expression is very similar to the reweighting factor necessary when initiating paths from a biased distribution of shooting points [14]. During the exchange, y acts as a shooting point to generate X' and y' , selected on X , is chosen with the same procedure as a shooting point in regular TPS. Therefore, we call the exchange scheme *shooting point exchange* (SPEX) in the following.

We first test the sampling scheme on a double well model, where stable states are connected by two distinct reaction channels (Figure 2A,B). In this system, two factors complicate the efficient sampling of configuration and path space. On the one hand, stable states are separated by an energy barrier limiting the occurrence of switches between them. On the other hand, also the two reaction channels are separated by a barrier. Hence, sampling transition paths connecting states A and B suffers from strong correlations since subsequently visited paths tend to remain in the same reaction channel. We compare the performance of standalone metadynamics and TPS with the performance of SPEX for different heights of the barrier separating the stable states (simulation details in supplementary information, SI [21]). In configuration space, sampled with metadynamics, we measure the time needed to switch between stable states A and B, $\tau_{\text{A} \leftrightarrow \text{B}}$. The exchange moves decrease the switching time between the two states at all barrier heights (Figure 2C). As the speedup is linked to the number of accepted exchanges, the effect is more pronounced at smaller barrier heights

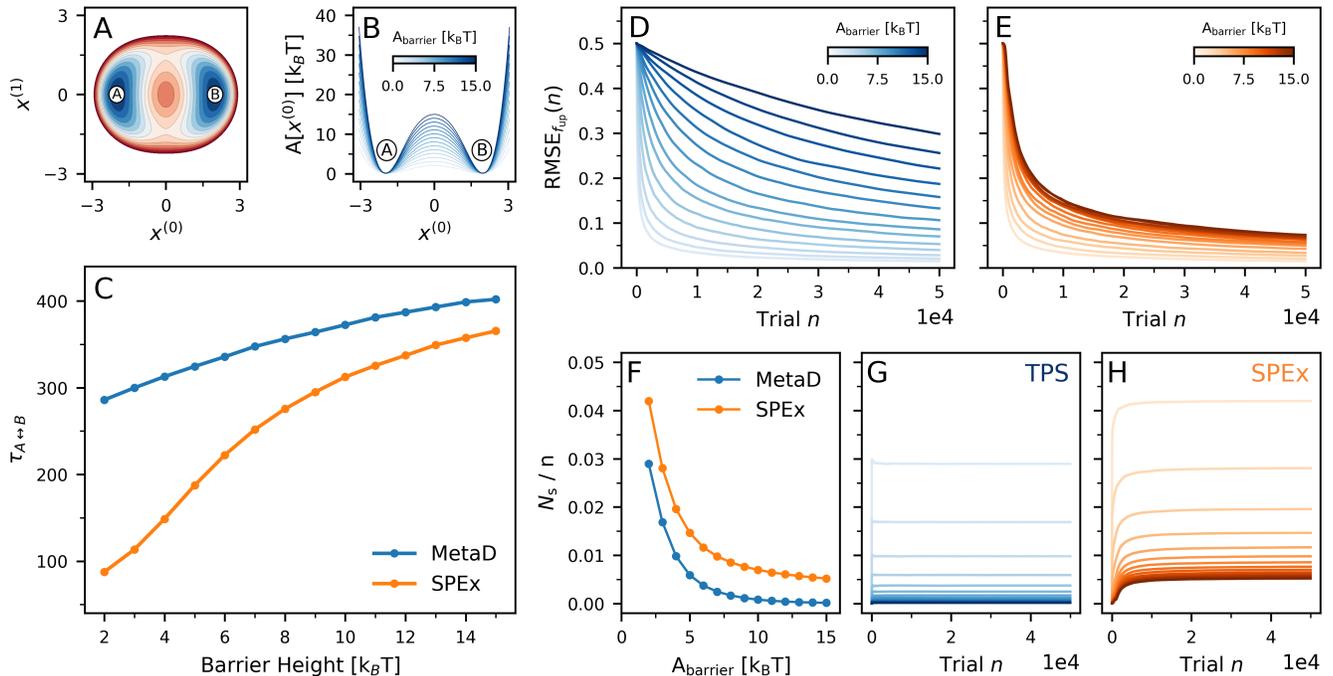


Figure 2. Shooting point exchange for sampling the two reaction channels in the two-dimensional double well model with two reaction channels. (A) Potential energy and state definitions of the model system. (B) Free energy as a function of $x^{(0)}$ for different barrier heights. (C) Average switching time to observe a switch between states A and B, $\tau_{A \leftrightarrow B}$, for standalone metadynamics and SPEx, starting the sampling from a converged bias potential. (D, E) Root mean squared error of the fraction of paths in the upper reaction channel as a function of simulation length for standalone TPS and SPEx. Each curve is for a specific barrier height and is estimated from 2500 independent sampling runs. (F, G, H) Number of switches between the upper and lower reaction channel N_s as a function of the barrier height (F) and the number of trials (G, H).

due to an increased likelihood to generate a transition path also from configurations away from the barrier.

For transition paths, a limiting factor for the sampling of trajectories that connect A and B is the slow switching between the upper and lower reaction channel. In regular TPS, the fraction of paths taking the upper reaction channel converges very slowly to the analytical fraction of $f_{\text{up}} = \frac{1}{2}$. This is apparent looking at the root mean square error of f_{up} as a function of the trial number n estimated from N runs (Figure 2D, E):

$$\text{RMSE}_{f_{\text{up}}}(n) = \sqrt{\sum_{i=0}^N \left(f_{\text{up}}(n) - \frac{1}{2} \right)^2} \quad (13)$$

In comparison, propagation of the system using metadynamics with shooting point exchange speeds up the convergence substantially (Figure 2E), especially for high barriers. This can be traced back to an increased number of switches during the sampling of the path ensemble (Figure 2F, G, H). Above barriers of $10 k_B T$, often not even a single switch between the channels occurs within 1000 trials in standalone TPS. When exchange moves are included, a minimum number of switches is recovered, mostly mediated by the exchange moves themselves.

From these observations, we can conclude that exchanges between an enhanced sampling simulation in

configuration space and a path sampling simulation increase the sampling efficiency on either side compared to running both simulations separately. While the increased switching time $\tau_{A \leftrightarrow B}$ is not negligible, the additional cost of sampling a path ensemble in parallel is not compensated. Therefore, at least in this setup, SPEx is not expected to increase the sampling efficiency when the interest is solely on configuration space. In contrast, when sampling path space, the additional force evaluations from metadynamics per TPS trial are insignificant compared to the number of force evaluations needed for the generation of a new path.

As a second case study, we investigate the *cis* to *trans* isomerization of the amino acid proline in the tetrapeptide KPTP [22] (Figure 3A). Proline isomerization plays an important role in protein folding [23, 24] and signaling in cells [25, 26], yet it only occurs on the timescale of seconds to minutes [27]. Due to the periodic nature of the relevant imide torsion angle $\omega^{\text{P}2}$, the transition from *cis* to *trans* and vice versa can take place via different reaction channels (Figure 3B). During the transition from $\pm 180^\circ$ (*trans*) to 0° (*cis*), the torsion angle can either cross over a barrier at -90° or 108° , referred to as *anti* and *syn* conformation [28] (dashed lines in figure 3D). Additionally, the imide nitrogen geometry, which is planar in the stable *cis* and *trans* states, is deformed out

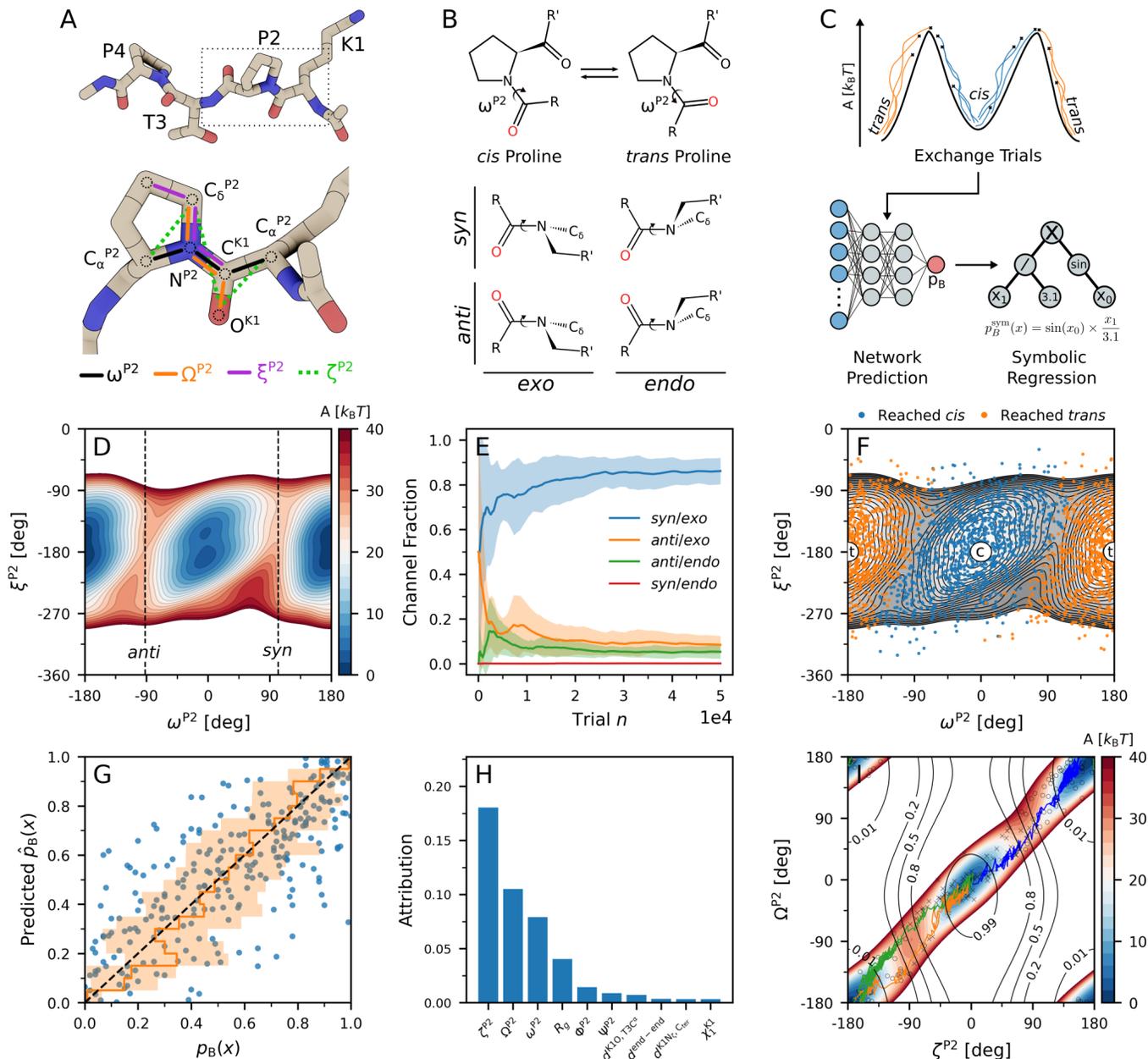


Figure 3. Isomerization of proline in the tetrapeptide KPTP. (A) Structure of the peptide (relevant torsion angle definitions in the inset) (B) Scheme of the *cis* to *trans* isomerization and the transition state geometry (C) Schematic overview of the committor learning process. (D) Free energy from metadynamics including shooting point exchanges. (E) Fraction of paths in each reaction channel as a function of the number of trials based on ten independent shooting point exchange simulations. (F) Training data for the committor prediction on top of the free energy surface. Circles show the state definitions (*c*=*cis*, *t*=*trans*). (G) Comparison of the sampled committor from fleeting trajectories and the predicted committor. The dashed black line shows the ideal correspondence while the orange line and shaded area show the average and standard deviation of the sampled committor in a given window of the predicted committor. (H) Attributions corresponding to the ten most important input features of the neural network. (I) Free energy along the two most relevant collective variables from (H). Black lines show isolines of the committor function obtained using symbolic regression. Crosses and circles indicate if a trajectory starting from that point reached the *cis* or *trans* state first. Representative reactive paths are shown in the same color scheme as in (E).

of plane [28]. By the direction of the deformation indicated by the torsion angle η_{P2} , the transition state can be distinguished as *endo* or *exo*, resulting in a total of four channels. The critical out of plane deformation is not

captured by ω^{P2} and therefore previous works proposed an improper dihedral ζ_{P2} as reaction coordinate [28, 29] (Figure 3A). Other collective variables discussed in the context of the isomerization are the Ψ_{P2} backbone an-

gle [28–30], the puckering state [31, 32] of the ring and solvent interactions [33].

Previous studies that focused on the mechanism of proline isomerization mainly used biased molecular dynamics to enhance the sampling [22, 29–32]. As a result, the dynamics of the system was altered and conclusions on the preferred mechanism and a corresponding reaction coordinate were mostly drawn based on minimum energy paths, with a notable exception being the recent work by Moritsugu et al. on the Pin1 enzyme [34]. We aim to identify the preferred isomerization mechanism, search systematically for relevant degrees of freedom and refine a reaction coordinate based on the unbiased dynamics of the system (simulation details in SI [21]). Besides the imide torsion angle, we choose ξ^{P2} (see Figure 3A) for sampling, as we expected it to capture potential geometric changes of both the imide nitrogen and the sidechain. The resulting free energy from metadynamics with shooting point exchanges agrees with previous studies in terms of the barrier height and difference between the *cis* and *trans* state [22, 30] (Figure 3D). However, estimating the different statistical weights of the four reaction channels is not possible. The *endo* and *exo* paths are not discriminated by ξ^{P2} and, more importantly, an estimation based on barrier heights does not account for entropy in path space. Here, the sampled transition paths can give an accurate estimate of the fraction of paths going through each channel (Figure 3E). These are *syn/exo* 0.862, *anti/exo* 0.084, *anti/endo* 0.053 and *syn/endo* 0.001, pointing out a clear preference for the *syn/exo* pathway.

Intending to find an improved reaction coordinate for the transition, we train a neural network to predict the committor probability $p_B(x)$ of a given configuration as proposed by Jung et al. [35]. The committor describes the likelihood of reaching state B before state A starting a simulation from configuration x and thereby also describes the progress of a reaction. From a broad set of collective variables based on which the network predicts the committor, the most important features can be determined by assigning an attribution score [35]. In the context of SPEX, we train the network using the information obtained from exchange moves (Figure 3F, network details in SI [21]). Although the training data only contain labels indicating if *cis* or *trans* was reached first, the network learns to interpolate in ambiguous regions (Figure 3G). Looking at the assigned attributions (Figure 3H), the torsion angles ζ^{P2} , Ω^{P2} and ω^{P2} are the most important variables, followed by the radius of gyra-

tion R_g , which has previously been discussed to be linked to the fraction of *cis* proline residues [22]. Collective variables describing the puckering state of the ring, Ψ_{P2} and all other backbone angles do not contribute significantly to the prediction of the committor. The neural network prediction is then used to refine an expression for a reaction coordinate via symbolic regression. Here, we include only the three torsion angles in the analysis in an attempt to obtain a reaction coordinate independent of the peptide sequence.

The most accurate estimate of $p_B(x)$ from symbolic regression includes ζ^{P2} and Ω^{P2} (Figure 3I):

$$p_B(\zeta^{P2}, \Omega^{P2}) = \text{sig}[-\sin(\Omega^{P2} - 0.75) + 4.334 \cos(\zeta^{P2}) + \cos(\Omega^{P2}) - 0.635], \quad (14)$$

where $\text{sig}(x) = 1/[1 + \exp(-x)]$. Although ζ^{P2} is undoubtedly a better reaction coordinate than ω^{P2} as shown in previous works, the committor isolines indicate that at least Ω^{P2} is required for an accurate prediction of $p_B(x)$.

To conclude, we presented a framework based on exchange moves between a configuration and a path ensemble. The sampling scheme has the potential to efficiently explore free energy surfaces, transition path ensembles and reaction coordinates of molecular processes, as demonstrated on the proline *cis-trans* isomerization. The case studies presented here — combining metadynamics and TPS — are just one realization of the possibilities emerging from Eq. (9). Since the generalized ensemble is not limited to a single configuration and path ensemble, we see future applications e.g. in umbrella sampling [8], multi-state TPS [36] or TIS [37].

Data Availability

The data that support the findings of this study are available upon reasonable request.

Acknowledgments

We acknowledge financial support of the Austrian Science Fund (FWF) through the SFB TACO, Grant number F 81-N. The computational results presented were achieved using the Vienna Scientific Cluster (VSC).

[1] Arjun, T. A. Berendsen, and P. G. Bolhuis, Proceedings of the National Academy of Sciences **116**, 19305 (2019).
 [2] G. Menzl and C. Dellago, Journal of Chemical Physics **145**, 10.1063/1.4964327 (2016).

[3] C. Leitold, C. J. Mundy, M. D. Baer, G. K. Schenter, and B. Peters, The Journal of Chemical Physics **153**, 024103 (2020).
 [4] J. Juraszek and P. G. Bolhuis, Proceedings of the National Academy of Sciences **103**, 15859 (2006).

- [5] K. ichi Okazaki, D. Wöhlert, J. Warnau, H. Jung, Özkan Yildiz, W. Kühlbrandt, and G. Hummer, *Nature Communications* **10**, 1742 (2019).
- [6] A. Laio and M. Parrinello, *Proceedings of the National Academy of Sciences* **99**, 12562 (2002).
- [7] A. Barducci, G. Bussi, and M. Parrinello, *Physical Review Letters* **100**, 020603 (2008).
- [8] G. Torrie and J. Valleau, *Journal of Computational Physics* **23**, 187 (1977).
- [9] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *The Journal of Chemical Physics* **108**, 1964 (1998).
- [10] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Annual Review of Physical Chemistry* **53**, 291 (2002).
- [11] P. G. Bolhuis and G. Csányi, *Physical Review Letters* **120**, 250601 (2018).
- [12] E. Borrero and C. Dellago, *The European Physical Journal Special Topics* **225**, 1609 (2016).
- [13] D. Mandelli, B. Hirshberg, and M. Parrinello, *Physical Review Letters* **125**, 026001 (2020).
- [14] S. Falkner, A. Coretti, S. Romano, P. Geissler, and C. Dellago, arXiv:2207.14530 (2022).
- [15] G. Hummer, *The Journal of Chemical Physics* **120**, 516 (2004).
- [16] R. H. Swendsen and J.-S. Wang, *Physical Review Letters* **57**, 2607 (1986).
- [17] Y. Sugita and Y. Okamoto, *Chemical Physics Letters* **314**, 141 (1999).
- [18] T. S. van Erp, *Physical Review Letters* **98**, 268301 (2007).
- [19] A. Gil-Ley and G. Bussi, *Journal of Chemical Theory and Computation* **11**, 1077 (2015).
- [20] H. Jung, K. ichi Okazaki, and G. Hummer, *The Journal of Chemical Physics* **147**, 152716 (2017).
- [21] See Supplemental Material at [URL will be inserted by publisher] for simulation and network details. The document contains refs. [38–43].
- [22] J. Alcantara, R. Stix, K. Huang, A. Connor, R. East, V. Jaramillo-Martinez, E. J. Stollar, and K. A. Ball, *Frontiers in Molecular Biosciences* **8** (2021).
- [23] W. J. Wedemeyer, E. Welker, and H. A. Scheraga, *Biochemistry* **41**, 14637 (2002).
- [24] F. Favretto, D. Flores, J. D. Baker, T. Strohäker, L. B. Andreas, L. J. Blair, S. Becker, and M. Zweckstetter, *Nature Communications* **11**, 6046 (2020).
- [25] K. P. Lu, G. Finn, T. H. Lee, and L. K. Nicholson, *Nature Chemical Biology* **3**, 619 (2007).
- [26] P. Sarkar, C. Reichman, T. Saleh, R. B. Birge, and C. G. Kalodimos, *Molecular Cell* **25**, 413 (2007).
- [27] C. Grathwohl and K. Wüthrich, *Biopolymers* **20**, 2623 (1981).
- [28] S. Fischer, R. L. Dunbrack, and M. Karplus, *Journal of the American Chemical Society* **116**, 11931 (1994).
- [29] C. Melis, G. Bussi, S. C. R. Lummis, and C. Molteni, *The Journal of Physical Chemistry B* **113**, 12148 (2009).
- [30] G. P. D. Martino, M. Masetti, A. Cavalli, and M. Recanatini, *Proteins: Structure, Function, and Bioinformatics* **82**, 2943 (2014).
- [31] Y. K. Kang and H. Y. Choi, *Biophysical Chemistry* **111**, 135 (2004).
- [32] D. Wu, *AIP Advances* **3**, 032141 (2013).
- [33] H. Ke, D. Mayrose, and W. Cao, *Proceedings of the National Academy of Sciences* **90**, 3324 (1993).
- [34] K. Moritsugu, N. Yamamoto, Y. Yonezawa, S. ichi Tate, and H. Fujisaki, *Journal of Chemical Theory and Computation* **17**, 2522 (2021).
- [35] H. Jung, R. Covino, A. Arjun, P. G. Bolhuis, and G. Hummer, arXiv:2105.06673 (2021).
- [36] J. Rogal and P. G. Bolhuis, *The Journal of Chemical Physics* **129**, 224107 (2008).
- [37] T. S. van Erp, D. Moroni, and P. G. Bolhuis, *The Journal of Chemical Physics* **118**, 7762 (2003).
- [38] N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, and H. J. C. Berendsen, *Journal of Chemical Theory and Computation* **8**, 3637 (2012).
- [39] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *The Journal of Chemical Physics* **79**, 926 (1983).
- [40] P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande, *PLOS Computational Biology* **13**, e1005659 (2017).
- [41] M. Bonomi, G. Bussi, C. Camilloni, G. A. Tribello, P. Banáš, A. Barducci, M. Bernetti, P. G. Bolhuis, S. Bottaro, D. Branduardi, R. Capelli, P. Carloni, M. Ceriotti, A. Cesari, H. Chen, W. Chen, F. Colizzi, S. De, M. D. L. Pierre, D. Donadio, V. Drobot, B. Ensing, A. L. Ferguson, M. Filizola, J. S. Fraser, H. Fu, P. Gasparotto, F. L. Gervasio, F. Giberti, A. Gil-Ley, T. Giorgino, G. T. Heller, G. M. Hocky, M. Iannuzzi, M. Invernizzi, K. E. Jelfs, A. Jussupow, E. Kirilin, A. Laio, V. Limongelli, K. Lindorff-Larsen, T. Löhner, F. Marinelli, L. Martin-Samos, M. Masetti, R. Meyer, A. Michaelides, C. Molteni, T. Morishita, M. Nava, C. Papissoni, E. Papaleo, M. Parrinello, J. Pfendtner, P. Piaggi, G. Piccini, A. Pietropaolo, F. Pietrucci, S. Pipolo, D. Provasi, D. Quigley, P. Raiteri, S. Raniolo, J. Rydzewski, M. Salvaglio, G. C. Sosso, V. Spiwok, J. Šponer, D. W. H. Swenson, P. Tiwary, O. Valsson, M. Vendruscolo, G. A. Voth, A. White, and T. P. consortium, *Nature Methods* **16**, 670 (2019).
- [42] D. A. Sivak, J. D. Chodera, and G. E. Crooks, *The Journal of Physical Chemistry B* **118**, 6466 (2014).
- [43] M. Huang, T. J. Giese, T.-S. Lee, and D. M. York, *Journal of Chemical Theory and Computation* **10**, 1538 (2014).

Supplementary Information: Enhanced Sampling of Configuration and Path Space in a Generalized Ensemble by Shooting Point Exchange

Sebastian Falkner,¹ Alessandro Coretti,¹ and Christoph Dellago^{1,*}

¹ *University of Vienna, Faculty of Physics, 1090 Vienna, Austria.*

*Electronic address: christoph.dellago@univie.ac.at

(Dated: March 23, 2023)

SI. Simulation Details for the Double Well System

The double well has the potential energy form:

$$U(x) = \alpha \left\{ 0.25 [(x^{(0)})^2 + (x^{(1)})^2 - 4]^2 + (x^{(1)})^2 \right\} \quad (\text{S1})$$

where α was adjusted to match the desired barrier height. All simulations were run using an underdamped Langevin integrator [S1] with a friction of 20, a timestep of 0.01, a mass of 1 and $k_{\text{B}}T = 1$. In simulations with and without shooting point exchange (SPEX), well-tempered metadynamics was configured identically placing a Gaussian of width 0.25 every 100 steps. The initial Gaussian height and bias factor were adjusted according to $0.2/15 \times A_{\text{barrier}}$ and $1 + (10/15 \times A_{\text{barrier}})$ respectively, where A_{barrier} is the barrier height. The current bias was stored on a grid with a bin width of 0.1 times the Gaussian width. Simulations with shooting point exchange followed the protocol of attempting an exchange every one TPS trial and 100 metadynamics steps.

SII. Simulation Details for the KPTP-Peptide

We prepared a simulation box with an edge length of 4.95 nm including the KPTP tetrapeptide solvated in TIP3P water [S2]. We added K^+ and Cl^- ions up to a concentration of 150 mmol L^{-1} to neutralize the box. All simulations were performed employing OpenMM [S3] and PLUMED [S4]. We used a velocity Verlet with velocity randomization integrator [S5] for simulation in the NVT ensemble at 310 K. The timestep was set to 2 fs and the friction to 1 ps^{-1} . All hydrogen bonds were constrained and the center of mass motion was removed at each timestep. Electrostatic interactions were treated using PME and the non-bonded cutoff was set to 1.2 nm.

For SPEX, a well-tempered metadynamics and transition path sampling simulation were run independently and exchanges between them were performed every one TPS trial and 2000 metadynamics simulation steps. For metadynamics, we bias along the ω^{P2} and ξ^{P2} as described in the main text with a bias factor of 20. The Gaussian width in both dimensions was 0.3 rad and the initial height was set to 2 kJ mol^{-1} . A Gaussian kernel was placed every 100 integration steps.

For the path sampling simulations, we define stable states as a function of ω^{P2} and ξ^{P2} :

$$h_{\text{trans}}(x) = 1 \text{ if } (\omega^{\text{P2}} - \pi)^2 + (\xi^{\text{P2}} - \pi)^2 < 0.0625 \quad (\text{S2})$$

$$h_{\text{cis}}(x) = 1 \text{ if } (\omega^{\text{P2}})^2 + (\xi^{\text{P2}} - \pi)^2 < 0.0625 \quad (\text{S3})$$

and set a maximum path length of 2 ns. Shooting points were selected based on the current bias potential as described in the main text and velocities were redrawn from a Maxwell-Boltzmann distribution after selection to decorrelate paths faster.

SIII. Training Protocol for Committor Learning

Each exchange trial provides data on the committor in form of a configuration x and whether the simulation reached *cis* or *trans* (0 / 1) from this point. We use a neural network (see table S1) to predict the expected outcome based

on this data. Instead of x , we provide a set of collective variables (see table S2) and the network output is restricted to the range $[0, 1]$. The collective variables are normalized by subtracting their mean and dividing by their respective standard deviations. Each torsion angle is then mapped on two scalars in form of the sine and cosine of its value to enforce periodicity of the network output. The training is performed in Pytorch using a loss function of the form:

$$L = \frac{1}{N} \sum_{i=0}^{N-1} \log[\epsilon + e^{-\hat{y}_i y_i}] \quad (\text{S4})$$

where N is the number of samples in the batch, \hat{y}_i is the predicted label and y_i the reference label. With $\epsilon = 1$, the loss function would be equivalent to a soft-margin loss, however, we use $\epsilon = 2.5$ to reduce the penalty of a misclassification. These are common since the training set only includes zeros or ones but we aim to predict probabilities that lie in between both values. We train for 15 epochs on 2.5×10^5 data points using the Adam optimizer with a learning rate of 0.005, a weight decay of 0.0001 and an exponential learning rate decay with a decay rate of 0.8. For improved accuracy, we train an ensemble of 10 models and average their committor prediction for the final result. Attributions (table S2) were assigned by feature permutation as described in Jung et al. [S6].

The symbolic regression was performed using gplearn on 2500 data points comprised of the three most important features from the attribution analysis and their network-predicted committor value. Parameters for reproducing the genetic search are given in table S3. The fitness function applies a sigmoid to all search results and calculates the mean squared error of the prediction with respect to the target from the neural network output. We start 100 independent optimization runs which we find more efficient in exploration than a larger population size or smaller tournament size. The ten best-performing models from these symbolic regression runs are shown in table S4.

-
- [S1] N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, and H. J. C. Berendsen, *Journal of Chemical Theory and Computation* **8**, 3637 (2012), ISSN 1549-9618.
- [S2] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *The Journal of Chemical Physics* **79**, 926 (1983), ISSN 0021-9606, URL <http://aip.scitation.org/doi/10.1063/1.445869>.
- [S3] P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, et al., *PLOS Computational Biology* **13**, e1005659 (2017), ISSN 1553-7358.
- [S4] M. Bonomi, G. Bussi, C. Camilloni, G. A. Tribello, P. Banáš, A. Barducci, M. Bernetti, P. G. Bolhuis, S. Bottaro, D. Branduardi, et al., *Nature Methods* **16**, 670 (2019), ISSN 1548-7105, URL <https://doi.org/10.1038/s41592-019-0506-8>.
- [S5] D. A. Sivak, J. D. Chodera, and G. E. Crooks, *The Journal of Physical Chemistry B* **118**, 6466 (2014), ISSN 1520-6106.
- [S6] H. Jung, R. Covino, A. Arjun, P. G. Bolhuis, and G. Hummer, arXiv:2105.06673 (2021), URL <http://arxiv.org/abs/2105.06673>.
- [S7] S. Fischer, R. L. Dunbrack, and M. Karplus, *Journal of the American Chemical Society* **116**, 11931 (1994), ISSN 0002-7863, URL <https://pubs.acs.org/doi/abs/10.1021/ja00105a036>.
- [S8] M. Huang, T. J. Giese, T.-S. Lee, and D. M. York, *Journal of Chemical Theory and Computation* **10**, 1538 (2014), ISSN 1549-9618.

Table S1. Neural network architecture for the prediction of the committor in the KPTP tetrapeptide system.

Layer	Components
Input Layer	Batch Norm Linear $N_{CV_s} \rightarrow 128$ Batch Norm + ReLU Activation
Layer 1	Linear $128 \rightarrow 64$ Batch Norm + ReLU Activation
Layer 2	Linear $64 \rightarrow 32$ Batch Norm + ReLU Activation
Layer 3	Linear $32 \rightarrow 16$ Batch Norm + ReLU Activation
Layer 4	Linear $16 \rightarrow 8$ Batch Norm + ReLU Activation
Output Layer	Linear $8 \rightarrow 1$ Sigmoid

Table S2. Collective variable descriptors and their attribution score for the neural network-based prediction of the committor. On the right, the features are grouped by type (e.g. distance, angle, ...) and on the left they are sorted by their absolute attribution score. If no reference is given, notations are as follows: d distance, θ angle, φ torsion angle, $[\Phi, \Psi, \omega, \chi]$ backbone/sidechain torsion angles.

Collective Variable	Attribution	Collective Variable	Attribution
$d^{\text{K1N}\zeta, \text{T3O}\gamma}$	2.85E-03	ζ^{P2} [S28]	1.80E-01
$d^{\text{K1O}, \text{K1N}}$	-1.10E-04	Ω^{P2}	1.05E-01
$d^{\text{K1N}\zeta, \text{C}_{\text{ter}}}$	3.38E-03	ω^{P2}	7.91E-02
$d^{\text{K1O}, \text{ACE-C}}$	-2.29E-05	R_g	4.05E-02
$d^{\text{K1O}, \text{ACE-O}}$	-6.92E-04	Φ^{P2}	1.43E-02
$d^{\text{K1O}, \text{K1N}}$	2.10E-04	Ψ^{P2}	8.76E-03
$d^{\text{K1O}, \text{K1C}\alpha}$	-3.44E-05	$d^{\text{K1O}, \text{T3C}\alpha}$	7.26E-03
$d^{\text{K1O}, \text{K1C}\beta}$	-3.66E-04	$d^{\text{end-end}}$	3.46E-03
$d^{\text{K1O}, \text{K1C}\gamma}$	3.24E-08	$d^{\text{K1N}\zeta, \text{C}_{\text{ter}}}$	3.38E-03
$d^{\text{K1O}, \text{K1C}\delta}$	5.13E-04	χ_1^{K1}	3.31E-03
$d^{\text{K1O}, \text{K1C}\epsilon}$	-5.08E-08	$d^{\text{K1N}\zeta, \text{T3O}\gamma}$	2.85E-03
$d^{\text{K1O}, \text{K1C}}$	-6.35E-05	$d^{\text{K1O}, \text{P2O}}$	1.97E-03
$d^{\text{K1O}, \text{P2N}}$	-4.64E-05	$\varphi^{\text{K1O}, \text{K1C}, \text{K1C}\alpha, \text{K1C}\beta}$	-1.61E-03
$d^{\text{K1O}, \text{P2C}\delta}$	-2.74E-06	Puckering Phase [S43]	-1.40E-03
$d^{\text{K1O}, \text{P2C}\gamma}$	2.46E-04	χ_4^{P2}	-1.23E-03
$d^{\text{K1O}, \text{P2C}\beta}$	-2.56E-04	χ_1^{P2}	7.53E-04
$d^{\text{K1O}, \text{P2C}\alpha}$	8.64E-07	η^{P2} [S28]	7.49E-04
$d^{\text{K1O}, \text{P2C}}$	1.27E-07	χ_4^{K1}	7.24E-04
$d^{\text{K1O}, \text{P2O}}$	1.97E-03	$d^{\text{K1O}, \text{ACE-O}}$	-6.92E-04
$d^{\text{K1O}, \text{T3N}}$	-2.79E-05	Φ^{T3}	6.39E-04
$d^{\text{K1O}, \text{T3C}\alpha}$	7.26E-03	$d^{\text{K1O}, \text{K1C}\delta}$	5.13E-04
$\theta^{\text{K1O}, \text{K1C}, \text{K1C}\alpha}$	-3.98E-06	Puckering Amplitude [S43]	-5.07E-04
$\theta^{\text{K1C}\alpha, \text{K1C}, \text{P2N}}$	1.81E-04	$\theta^{\text{K1C}, \text{P2N}, \text{P2C}\alpha, \text{P2C}\beta}$	-4.84E-04
$\theta^{\text{K1C}, \text{P2N}, \text{P2C}\alpha, \text{P2C}\beta}$	-4.84E-04	ω^{T3}	-4.07E-04
$\varphi^{\text{K1O}, \text{K1C}, \text{K1C}\alpha, \text{K1N}}$	-1.06E-04	$d^{\text{K1O}, \text{K1C}\beta}$	-3.66E-04
$\varphi^{\text{K1O}, \text{K1C}, \text{K1C}\alpha, \text{K1C}\beta}$	-1.61E-03	ω^{P4}	3.47E-04
Ω^{P2}	1.05E-01	Puckering Z_y [S43]	-2.60E-04
$\varphi_{\alpha}^{\text{K1O}, \text{K1C}, \text{P2N}, \text{P2C}}$	9.16E-08	$d^{\text{K1O}, \text{P2C}\beta}$	-2.56E-04
φ^{ACE}	-5.50E-05	$d^{\text{K1O}, \text{P2C}\gamma}$	2.46E-04
ω^{K1}	-3.63E-05	Ψ^{T3}	2.40E-04
Ψ^{K1}	9.98E-05	χ_2^{P2}	2.27E-04
Φ^{K1}	1.32E-04	$d^{\text{K1O}, \text{K1N}}$	2.10E-04
χ_1^{K1}	3.31E-03	$\theta^{\text{K1C}\alpha, \text{K1C}, \text{P2N}}$	1.81E-04
χ_2^{K1}	-8.33E-05	ξ^{P2}	-1.61E-04
χ_3^{K1}	-1.48E-04	χ_4^{P4}	-1.51E-04
χ_4^{K1}	7.24E-04	χ_3^{K1}	-1.48E-04
ω^{P2}	7.91E-02	Φ^{K1}	1.32E-04
Ψ^{P2}	8.76E-03	χ_1^{T3}	1.26E-04
Φ^{P2}	1.43E-02	$d^{\text{K1O}, \text{K1N}}$	-1.10E-04
χ_1^{P2}	7.53E-04	$\varphi^{\text{K1O}, \text{K1C}, \text{K1C}\alpha, \text{K1N}}$	-1.06E-04
χ_2^{P2}	2.27E-04	Ψ^{K1}	9.98E-05
χ_3^{P2}	6.49E-06	φ^{NME}	-9.48E-05
χ_4^{P2}	-1.23E-03	χ_2^{K1}	-8.33E-05
ξ^{P2}	-1.61E-04	Φ^{P4}	-7.67E-05
ω^{T3}	-4.07E-04	Ψ^{P4}	-6.91E-05
Ψ^{T3}	2.40E-04	$d^{\text{K1O}, \text{K1C}}$	-6.35E-05
Φ^{T3}	6.39E-04	χ_2^{P4}	-5.63E-05
χ_1^{T3}	1.26E-04	φ^{ACE}	-5.50E-05
ω^{P4}	3.47E-04	$d^{\text{K1O}, \text{P2N}}$	-4.64E-05
Ψ^{P4}	-6.91E-05	χ_1^{P4}	-3.93E-05
Φ^{P4}	-7.67E-05	ω^{K1}	-3.63E-05
χ_1^{P4}	-3.93E-05	$d^{\text{K1O}, \text{K1C}\alpha}$	-3.44E-05
χ_2^{P4}	-5.63E-05	$d^{\text{K1O}, \text{T3N}}$	-2.79E-05
χ_3^{P4}	-9.88E-06	$d^{\text{K1O}, \text{ACE-C}}$	-2.29E-05
χ_4^{P4}	-1.51E-04	χ_3^{P4}	-9.88E-06
φ^{NME}	-9.48E-05	χ_3^{P2}	6.49E-06
ζ^{P2} [S28]	1.80E-01	$\theta^{\text{K1O}, \text{K1C}, \text{K1C}\alpha}$	-3.98E-06
η^{P2} [S28]	7.49E-04	$d^{\text{K1O}, \text{P2C}\delta}$	-2.74E-06
Puckering Phase [S43]	-1.40E-03	$d^{\text{K1O}, \text{P2C}\alpha}$	8.64E-07
Puckering Amplitude [S43]	-5.07E-04	Puckering Z_x [S43]	2.54E-07
Puckering Z_x [S43]	2.54E-07	$d^{\text{K1O}, \text{P2C}}$	1.27E-07
Puckering Z_y [S43]	-2.60E-04	$\varphi_{\alpha}^{\text{K1O}, \text{K1C}, \text{P2N}, \text{P2C}}$	9.16E-08
N_{H} 0.5 nm around K1O	-5.13E-08	N_{H} 0.5 nm around K1O	-5.13E-08
R_g	4.05E-02	$d^{\text{K1O}, \text{K1C}\epsilon}$	-5.08E-08
$d^{\text{end-end}}$	3.46E-03	$d^{\text{K1O}, \text{K1C}\gamma}$	3.24E-08

Table S3. Parameters for the symbolic regression of the committor function.

Parameter	Value
Population Size	1000
Tournament Size	20
Constant Range	$-2\pi - 2\pi$
Generations	100
$P_{\text{crossover}}$	0.7
$P_{\text{subtree-mutation}}$	0.1
$P_{\text{hoist-mutation}}$	0.05
$P_{\text{point-mutation}}$	0.1
Parsimony Coefficient	0.0001
Function Set	Add, Sub., Div., Mul., Sin, Cos

Table S4. Symbolic regression results for the committor prediction with the lowest error.

Symbolic Regression Result	Mean Square Error
$-\sin(\Omega^{P^2} - 0.75) + 4.334 \cos(\zeta^{P^2}) + \cos(\Omega^{P^2}) - 0.635$	3.201E-02
$-1.62(\zeta^{P^2})^2 + \cos(\Omega^{P^2}) + \cos(\Omega^{P^2} - 5.266) + 3.354$	3.205E-02
$\sin(\Omega^{P^2} + 2.474) + 4.703 \cos(\zeta^{P^2}) + \cos(\Omega^{P^2}) - 0.743$	3.240E-02
$\sin(\Omega^{P^2} + 2.497) + 4.778 \cos(\zeta^{P^2}) + \cos(\cos(\Omega^{P^2}) + 4.099)$	3.249E-02
$3.5308 \cos(\Omega^{P^2}) + 2 \cos(\omega^{P^2}) + \cos(\zeta^{P^2} - 5.354) - 0.283$	3.255E-02
$4.706 \sin(\zeta^{P^2} - 4.62) + \cos(\Omega^{P^2}) + \cos(\Omega^{P^2} + 0.466) - 0.682$	3.260E-02
$\sin(\Omega^{P^2} + 2.387) + 4.658 \cos(\zeta^{P^2}) + \cos(\Omega^{P^2}) - 0.74$	3.277E-02
$-(\zeta^{P^2})^2 - \sin(\Omega^{P^2}) + \cos(\zeta^{P^2}) + 2 \cos(\Omega^{P^2}) + 1.918$	3.281E-02
$-0.889 \sin(\Omega^{P^2}) + 3.458 \cos(\Omega^{P^2}) + 2.186 \cos(\omega^{P^2})$	3.304E-02
$-(\Omega^{P^2})^2 + \sin(\zeta^{P^2} - 3.792) + \cos(\zeta^{P^2}) + \cos(\omega^{P^2}) + \cos(\cos(\omega^{P^2})) + 1.366$	3.319E-02