

Supporting Information

DeepPath detailed implementation

Notations

$\{\dots\}_{\text{size}}$ represents a list of arrays. If the size is omitted, it remains the same as last defined. All densely connected layers (Linear) have 256 neurons unless otherwise specified. $\{RC\}$ is a list of structures along the generated path in reduced-coordinate representation. $\{x^{\text{aa}}\}$ is a such list in Cartesian coordinates. They can also be represented in backbone frames ($\{T\}$) and sidechain dihedral angles components ($\{\alpha^{\text{cos}}, \alpha^{\text{sin}}\}$).

Explorer

Algorithm S1 Explorer

```
1: function EXPLORER( $\{RC^{\text{endpoint}}\}$ ,  $n_{\text{conv}}$ ,  $n_{\text{latent}} = 8$ )
2:    $\{s_0\} \leftarrow \{\emptyset\}_2$ 
3:   for  $i = 0$  to 1 do ▷ Shared weights
4:      $s_0[i] \leftarrow \text{relu}(\text{Linear}_{128}(RC_i^{\text{endpoint}}))$ 
5:   end for
6:    $z \leftarrow \text{LayerNorm}(\text{RandomNormal}())$ 
7:   for  $i = 1$  to  $n_{\text{latent}}$  do
8:      $z \leftarrow \text{LeakyReLU}_{0.2}(\text{Linear}_{128}(z))$ 
9:   end for
10:
11:    $\{s\} \leftarrow \{s_0\}$ 
12:   for  $i = 1$  to  $n_{\text{conv}}$  do
13:      $\{s\} \leftarrow \text{Repeat}(\{s\}, 2)$  ▷ Upsample
14:      $s \leftarrow s \cdot \text{LeakyReLU}_{0.2}(\text{Linear}_{128}(z))$  ▷ Modulate
15:      $s \leftarrow \text{Concat}(s_0[0], s, s_0[1])$  ▷ Pad with endpoint representations
16:      $s \leftarrow \text{LeakyReLU}_{0.2}(\text{Conv1D}_3(s))$  ▷ Convolution
17:      $s \leftarrow s / (\text{stddev}(s) + \epsilon)$  ▷ Normalization
18:   end for
19:    $\{RC\} \leftarrow \text{softplus}(\text{Conv1D}_1(s))$ 
20:    $\{RC\} \leftarrow \text{Concat}(RC_0^{\text{endpoint}}, \{RC\}, RC_1^{\text{endpoint}})$  ▷ Pad with endpoints
21:   return  $\{RC\}$ 
22: end function
```

Structure Builder

Algorithm S2 Structure Builder

```
1: function STRUCTUREBUILDER( $RC$ ,  $n_{\text{layer}} = 5$ )
2:    $s_0 \leftarrow \text{relu}(\text{Linear}_{512}(RC))$ 
3:    $s \leftarrow \text{Linear}(s_0)$ 
4:    $\{T_0\} \leftarrow \{\mathbb{1}_4\}_{\text{nres}}$  ▷ Initialize  $T$  as identity matrices
5:   for  $i = 1$  to  $n_{\text{layer}}$  do
6:      $s, \{\Delta T\}_{\text{nres}}, \{\alpha_i^{\cos}, \alpha_i^{\sin}\}_{\text{ndih}} \leftarrow \text{STRUCTUREBLOCK}(s, s_0)$ 
7:      $T_i \leftarrow T_{i-1} \cdot \Delta T$  ▷ Apply incremental transformation
8:   end for
9:    $\{T^{\text{aa}}\}_{\text{natom}}, \{x^{\text{aa}}\}_{\text{natom}} \leftarrow \text{ALLATOMCOORDINATES}(\{T_{n_{\text{layer}}}\}, \{\alpha_{n_{\text{layer}}}^{\cos}, \alpha_{n_{\text{layer}}}^{\sin}\})$ 
10:  return  $\{T^{\text{aa}}\}, \{x^{\text{aa}}\}, \{T_i\}, \{\alpha_i^{\cos}, \alpha_i^{\sin}\} \quad \forall i \in [1, n_{\text{layer}}]$ 
11: end function
```

Refer to Algorithm 1 in the Main Text for the StructureBlock().

Algorithm S3 All-atom coordinate constructor

```
1: function ALLATOMCOORDINATES(  $\{T\}_{\text{nres}}$ ,  $\{\alpha^{\text{cos}}, \alpha^{\text{sin}}\}_{\text{ndih}}$ ,  $\{\text{dih\_type}\}_{\text{ndih}}$  )
2:    $(\alpha^{\text{cos}}, \alpha^{\text{sin}}) \leftarrow (\alpha^{\text{cos}}, \alpha^{\text{sin}}) / \|\alpha\|$   $\triangleright$  Normalize  $\alpha^{\text{cos}}$  and  $\alpha^{\text{sin}}$ 
3:    $T^\alpha \leftarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha^{\text{cos}} & -\alpha^{\text{sin}} & 0 \\ 0 & \alpha^{\text{sin}} & \alpha^{\text{cos}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$   $\triangleright$  Construct  $T$  base on predicted rotation angles

4:    $\{T^{\text{dih}}\} \leftarrow \{\emptyset\}_{\text{ndih}}$ 
5:   for  $i = 1$  to  $n_{\text{dih}}$  do
6:     if  $\text{dih\_type}[i]$  is ‘ $\psi$ ’ or ‘ $\chi_1$ ’ then
7:        $T^{\text{base}} \leftarrow T[\text{res}[i]]$   $\triangleright$  Use backbone as base
8:     else
9:        $T^{\text{base}} \leftarrow T^{\text{dih}}[i - 1]$   $\triangleright$  Use the preceding sidechain fragment as base
10:    end if
11:     $T^{\text{dih}}[i] \leftarrow T^{\text{base}} \cdot T^{\text{lit}}[\text{dih\_type}[i]] \cdot T^\alpha[i]$ 
12:     $\triangleright$  Combine base and local transformations
13:  end for
14:
15:   $\{x^{\text{aa}}\} \leftarrow \{\vec{0}\}_{\text{natom}}$ 
16:  for  $i = 1$  to  $n_{\text{natom}}$  do
17:     $T^{\text{aa}} \leftarrow T^{\text{dih}}[\text{dih\_from\_atomID}[i]]$ 
18:     $x^{\text{aa}}[i] \leftarrow T^{\text{aa}} \cdot x^{\text{lit}}[i]$   $\triangleright$  Apply corresponding  $T^{\text{dih}}$  to atom coordinates
19:  end for
20:
21:  return  $\{T^{\text{aa}}\}, \{x^{\text{aa}}\}$ 
22: end function
```

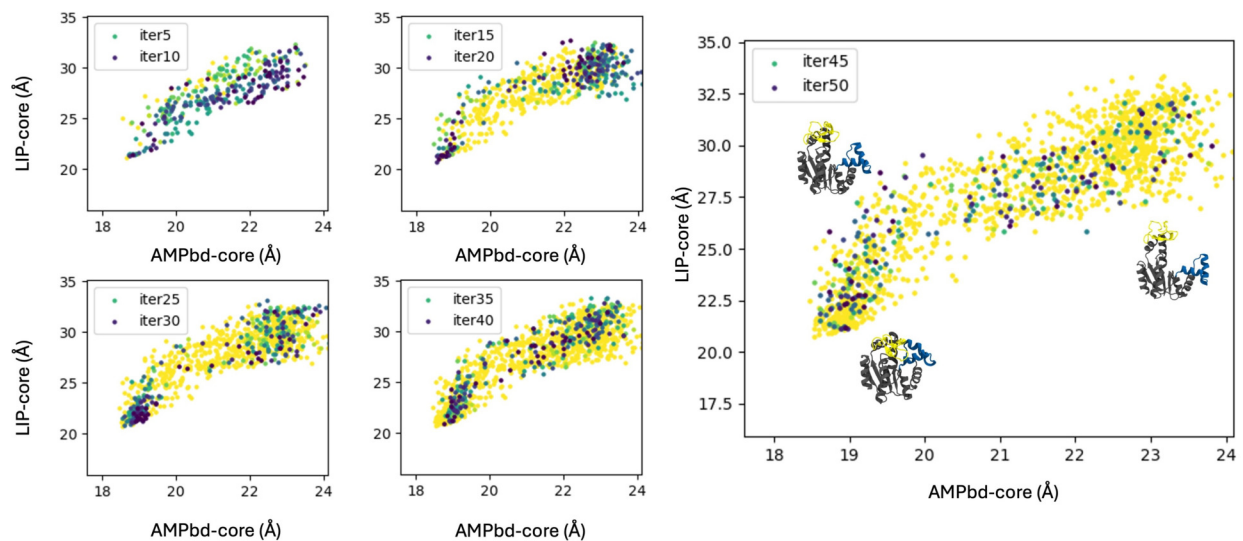


Figure S1: Iterative filling of the transition manifold for the ADK open-to-closed transition.

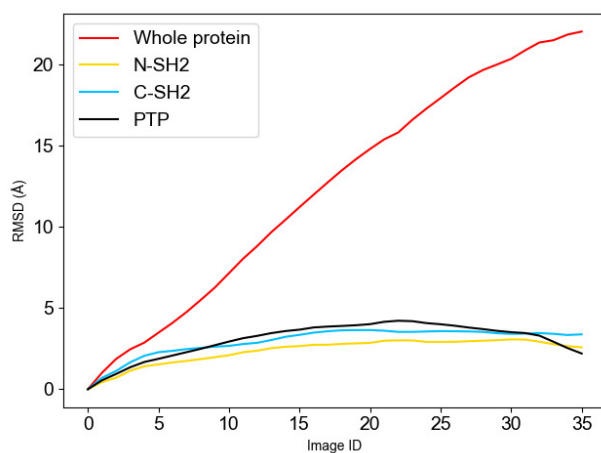


Figure S2: RMSD of the full SHP2 protein and its individual domains, computed relative to the predicted inactive state (image ID 0), for all structures along the predicted transition pathway. Only heavy atoms were used in the RMSD measurement.

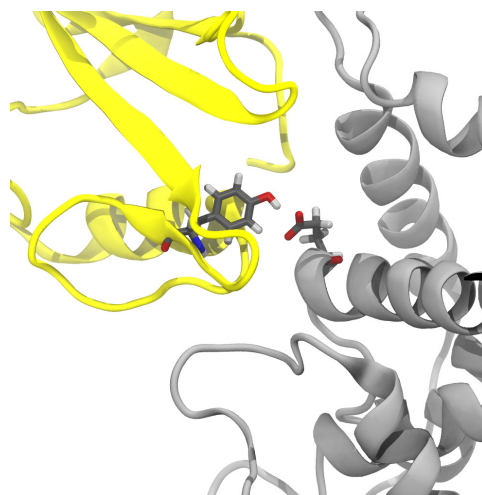


Figure S3: DeepPath identified the critical contact between residues Y63 and E508 of SHP2.

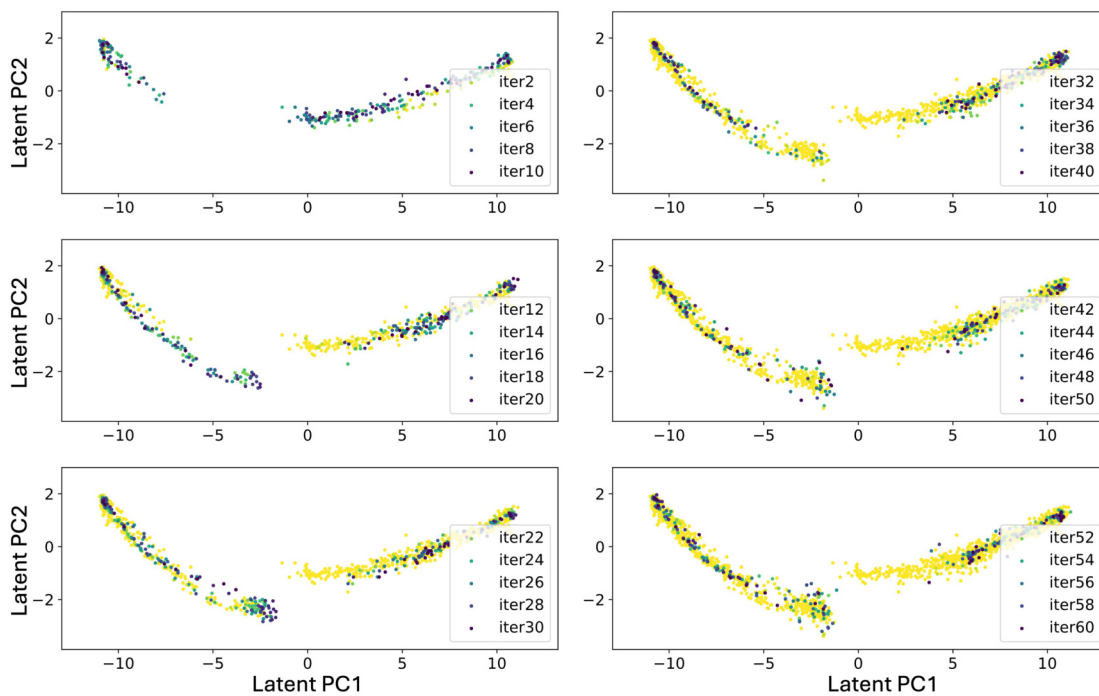


Figure S4: Latent space projections of DeepPath-generated structures during training using the AMBER force field. Each panel shows PCA projections (PC1 vs PC2) of the pair-distance representation, with colors indicating different training iterations.

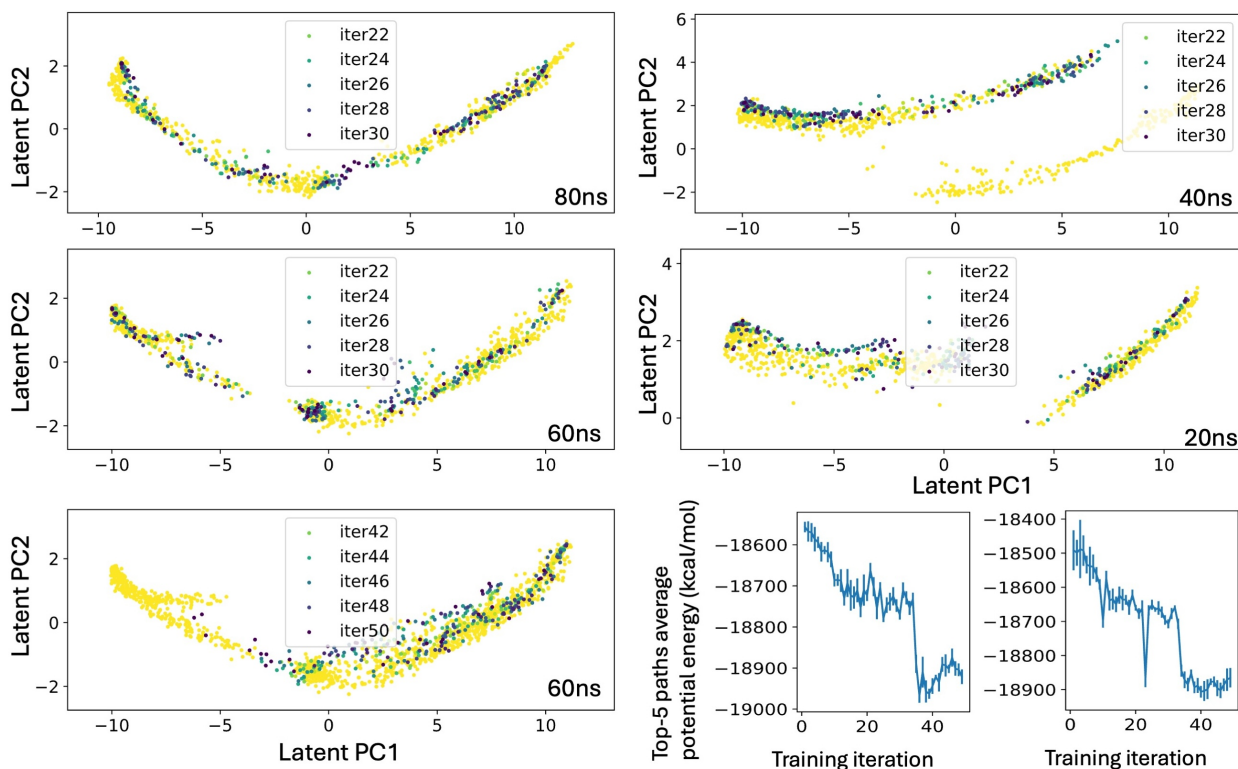


Figure S5: Effect of the length of the initial MD trajectory on DeepPath training for the SHP2 system. Generated structures are projected onto the same latent PCA space defined by the 100 ns dataset to enable direct comparison. Results are shown for training datasets derived from 80 ns (top left), 40 ns (top right), 60 ns (middle left), and 20 ns (middle right) MD trajectories, with points colored by training iteration. The shorter trajectories (20 ns and 40 ns) do not recover a continuous transition manifold, whereas the 60 ns and 80 ns datasets produce clearer transition pathways. The bottom-left panel shows later training iterations for the 60 ns dataset (iterations 40–50), illustrating the progressive filling of the transition manifold. The bottom-right panels show the evolution of the average potential energy of the top five predicted pathways during training for the 80 ns (left) and 60 ns (right) datasets, where the energy decreases as the model discovers lower-energy intermediates along the transition pathway.

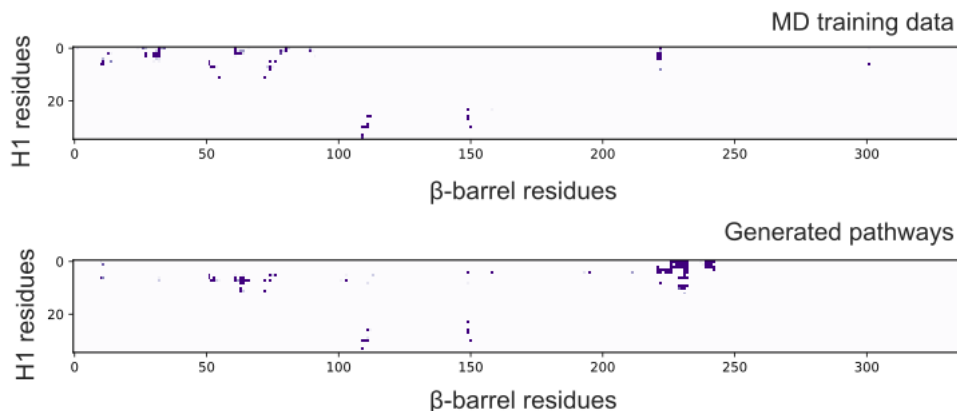


Figure S6: Contact maps between C- α atoms of the CdiB β -barrel and H1 helix using 7 Å as cutoff distance, identical the cutoff used for reduced coordinate construction. Residue pairs that are colored in the upper panel are include as part of the reduced coordinate set used by DeepPath. The lower panel shows the contacts exhibited by generated pathways, which includes contacts that are not part of the reduced coordinate set.

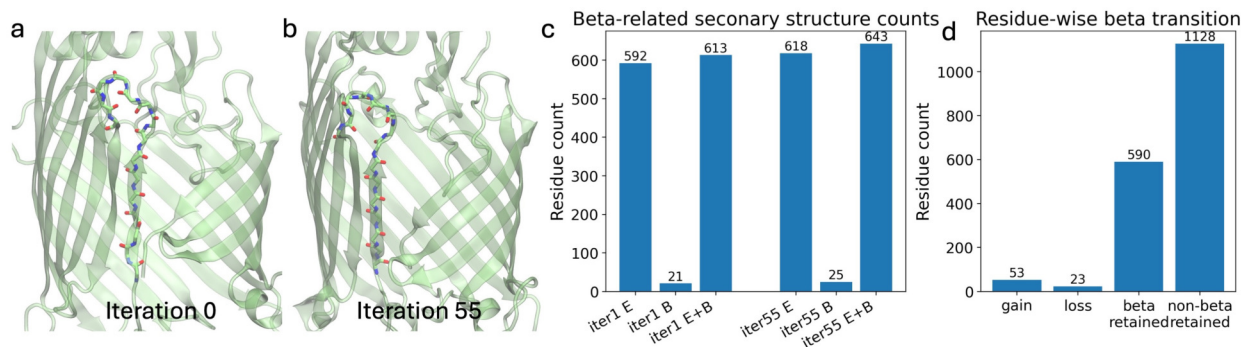


Figure S7: Correction of the BAM lateral gate secondary structure during DeepPath active learning. (a,b) Representative structures from early (iteration 1) and late (iteration 55) stages, showing the refinement of the lateral gate region toward a β -strand conformation. (c) Quantification of β -related secondary structure elements, showing an increase in β -content and a reduction of non- β assignments over training. (d) Residue-wise summary of secondary structure transitions, indicating that most residues either retain or adopt β -strand character, with relatively few losses.

Table S1: Energy statistics of CdiB path clusters found by DeepPath.

Cluster ID	No. of paths	Total energy (kcal/mol)	Interaction energy (kcal/mol)
0	52	-12072 \pm 6	-128 \pm 5
1	47	-12076 \pm 6	-129 \pm 7
2	59	-12076 \pm 6	-136 \pm 6
3	98	-12077 \pm 7	-139 \pm 5

Table S2: Inputs and key parameters used in DeepPath modeling. All notations follow the definitions in the Methods section.

Type	Item	Description	Value
Inputs			
Input	Protein force field	Protein force field used for energy evaluation in OpenMM	CHARMM36m (AdK, SHP2, BAM), CHARMM36 (CdiB)
Input	Solvent force field	Implicit solvent force field used for energy evaluation in OpenMM	GB ^{OBC} II (AdK, SHP2, CdiB), None (BAM)
Input	MD trajectories	Initial MD trajectories per end-state used to construct reduced coordinates and training data	1 ns (AdK), 10 ns (CdiB), 100 ns (SHP2), 500 ns (BAM)
Reduced coordinate definition			
Parameter	ϵ_A	Distance cutoff for defining residue contacts	6 Å (SHP2, BAM), 7 Å (CdiB)
Parameter	ϵ_B	Distance cutoff for defining non-contacts	15 Å (all systems)
Explorer generator loss components (Eq. 1)			
Loss	C_{WGAN}	WGAN critic score	weight = 1.0

Type	Item	Description	Value
Loss	$\mathcal{C}_{\text{energy}}$	Energy critic score	weight = 1.0
Loss	$\mathcal{C}_{\text{path}}$	Path continuity loss based on deviation from linear interpolation	weight = 10.0 (SHP2, BAM), weight = 2.0 (AdK, CdiB)
Loss	Clash penalty	Triangular-well potential penalizing distances below 3.5 Å	weight = 1.0
Loss	Diversity penalty	Linear penalty for structures with MAD below 0.5 Å in reduced coordinates	weight = 1.0

Explorer critics loss components

Loss	$\mathcal{L}_{\text{WGAN}}$	WGAN critic loss	weight = 1.0
Loss	\mathcal{L}_{GP}	Gradient penalty	weight = 10.0
Loss	$\mathcal{L}_{\text{zero-mean}}$	zero-centered score regularization	weight = 1.0
Loss	\mathcal{L}_{ES}	Energy critic student network loss	$\alpha = 0.1, T = 3$

Structure Builder loss functions

Loss	$\mathcal{L}_i^{\text{FAPE}}$	Frame-aligned point error for intermediate backbone frames	weight = 0.2
Loss	$\mathcal{L}_{\text{aa}}^{\text{FAPE}}$	FAPE of final all-atom structure	weight = 1.0
Loss	$\mathcal{L}_i^{\text{sidechain}}$	MSE of sidechain dihedral angles	weight = 0.2
Loss	$\mathcal{L}_{\text{bond}}$	Flat-bottom L1 loss enforcing amide bond length	weight = 0 (initial), weight = 1.0 (active learning)
Loss	$\mathcal{L}_{\text{clash}}$	Steric clash penalty for atoms closer than 1.2 Å	weight = 0 (initial), weight = 1.0 (active learning)

Type	Item	Description	Value
Loss	$\mathcal{L}_{\text{RMSD}}$	Sigmoid RMSD restraint on final structure	weight = 0.01 (BAM during active learning), weight = 0 (otherwise)

Training parameters

Parameter	Learning rate	Adam optimizer learning rate	0.001 (initial), 0.0001 (active learning)
Parameter	GAL iterations	Number of active learning cycles	50 (AdK, SHP2, CdiB), 55 (BAM)
Parameter	Selected structures	Number of structures added per iteration	32 (AdK, SHP2, CdiB), 24 (BAM)