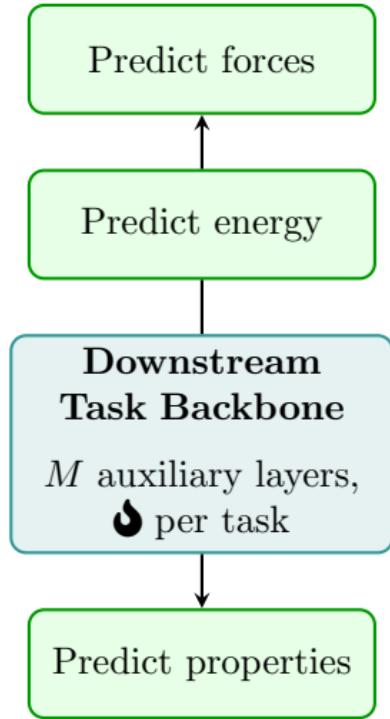


Downstream task heads



Atom types, coordinates,
lattice, time, etc.

**Input Embedding
& Positional Encoding**

Transformer Trunk
 L layers,
masked self-attention

[Finetuning]
Using specified
layer representation
✳

[Pretraining]
Using final layer
representation
🔥

Generative heads

Predict atom types

Predict coordinates

Predict lattice