

Supplementary Information

Inverse design of semiconductor materials with deep generative models

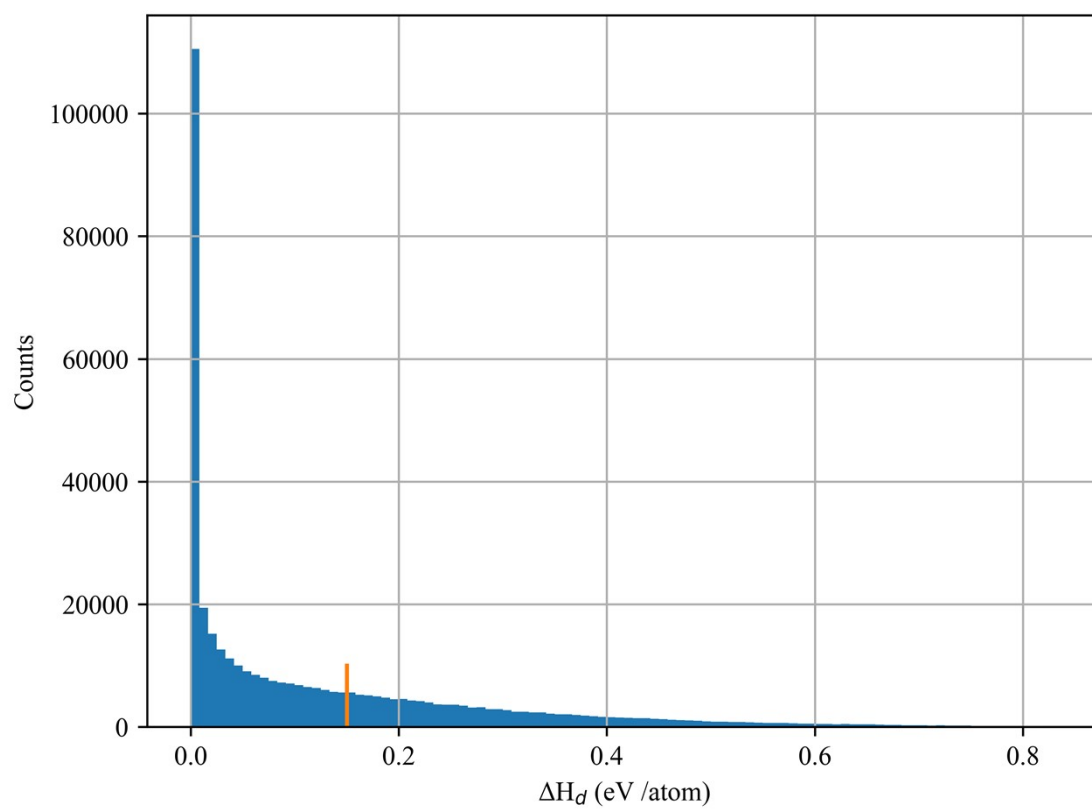


Fig. S1. Histogram of the distribution of decomposition enthalpy (ΔH_d) for the entire dataset. Approximately 2/3 of the samples in the dataset have a ΔH_d less than or equal to 0.15 eV/atom.

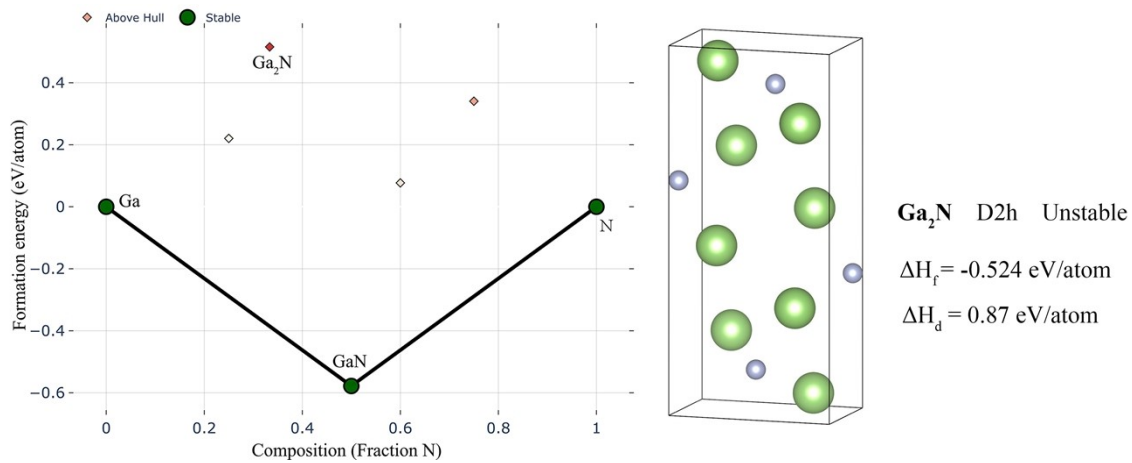


Fig. S2. The binary convex hull phase diagram for Ga-N systems incorporates the energies of Ga-N compositions from the dataset as well as the energy of Ga₂N predicted by VGD-CG.

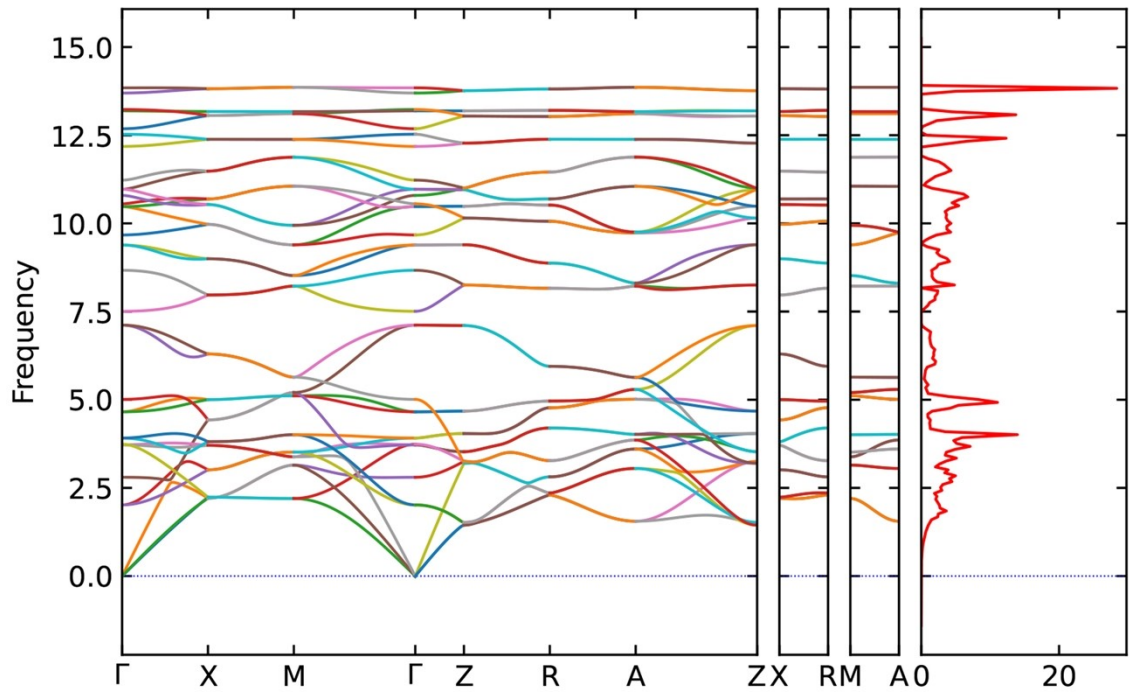


Fig. S3. The phonon dispersion curves calculated by PBE along symmetry lines in the Brillouin zone for Si_3Ge_2 .

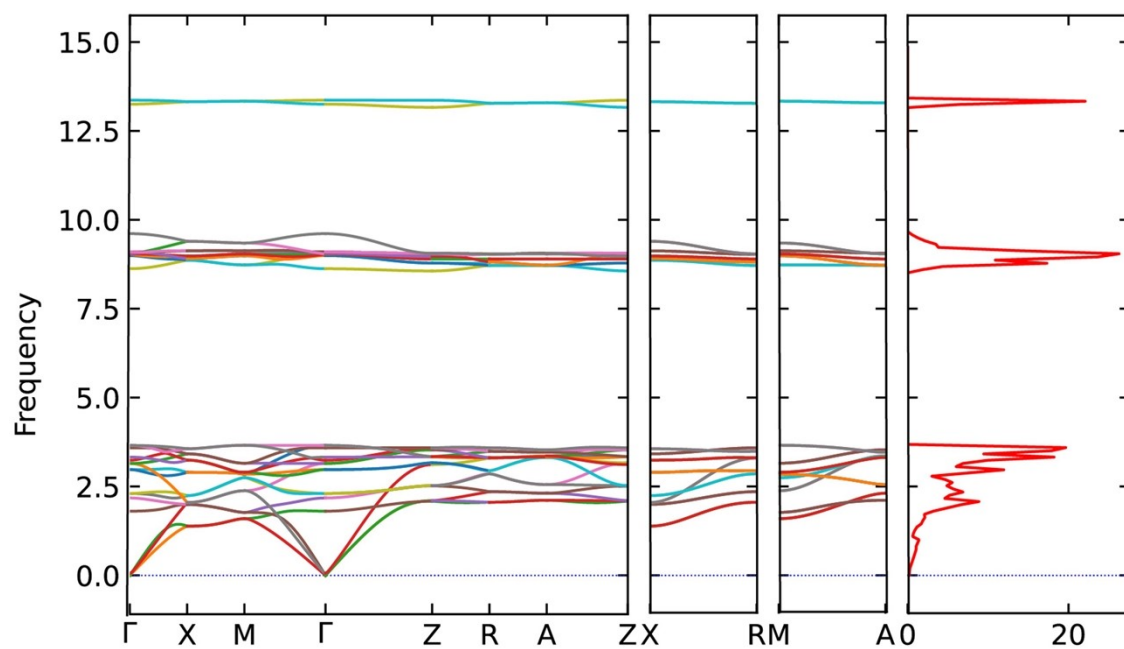


Fig. S4. The phonon dispersion curves calculated by PBE along symmetry lines in the Brillouin zone for VBi_3O_7 .

Table S1. The setup of model hyperparameters used in VAE, GAN and DDPM.

| Parameters | VAE | GAN | DDPM |
|----------------|-------|-------|----------------------|
| z_dim | 128 | 128 | \ |
| w_kl | 0.01 | \ | \ |
| batch_size | 5120 | 5120 | 256 |
| epochs | 20 | 20 | 20 |
| lr | 0.001 | 0.001 | 0.0001 |
| optimizer | Adam | Adam | Adam |
| train_ratio | 0.9 | 0.9 | 0.9 |
| test_ratio | 0.1 | 0.1 | 0.1 |
| n_critic | \ | 5 | \ |
| lambda_penalty | \ | 10 | \ |
| beta_end | \ | \ | 0.02 |
| timesteps | \ | \ | 500 |
| beta_start | \ | \ | 0.0001 |
| schedule_name | \ | \ | linear_beta_schedule |