| Property                          | Formula | Description                        | Fr. Name           |
|----------------------------------|---------|------------------------------------|--------------------|
| MaxEStateIndex                   | PEOE_VSA9 | NumAliphaticHeterocycles           | fr_benzodiazepine  |
| MinEStateIndex                   | SMR_VSA1 | NumAliphaticRings                  | fr_bicyclic       |
| MaxAbsEStateIndex                | SMR_VSA10 | NumAromaticCarbocycles             | fr_diazo           |
| MinAbsEStateIndex                | SMR_VSA2 | NumAromaticHeterocycles            | fr_dihydropyridine|
| qed                              | SMR_VSA3 | NumAromaticRings                   | fr_epoxide         |
| MolWt                            | SMR_VSA4 | NumHAcceptors                      | fr_esteer          |
| HeavyAtomMolWt                   | SMR_VSA5 | NumHDonors                         | fr_ether           |
| ExactMolWt                       | SMR_VSA6 | NumHeteroatoms                     | fr_furan           |
| NumValenceElectrons              | SMR_VSA7 | NumRotatableBonds                  | fr_guanido         |
| NumRadicalElectrons              | SMR_VSA8 | NumSaturatedCarbocycles            | fr_halogen         |
| MaxPartialCharge                 | SlogP_VSA1 | NumSaturatedRings                  | fr_hdrzone         |
| MinPartialCharge                 | SlogP_VSA10 | RingCount                          | fr_imidazole       |
| MaxAbsPartialCharge              | SlogP_VSA11 | MolLogP                            | fr_imide           |
| MinAbsPartialCharge              | SlogP_VSA12 | MolMR                              | fr_isocyan         |
| FpDensityMorgan1                 | SlogP_VSA2 | fr_Al_COO                          | fr_isothiocyan     |
| FpDensityMorgan2                 | SlogP_VSA3 | fr_Al_OH                           | fr_ketone          |
| FpDensityMorgan3                 | SlogP_VSA4 | fr_Al_OH_noTert                    | fr_ketone_TopliSS  |
| BalabanJ                         | SlogP_VSA5 | fr_ArN                             | fr_lactam          |
| Kappa1                           | SlogP_VSA6 | fr_Ar_COO                          | fr_lactone         |
| Chi0                             | SlogP_VSA7 | fr_Ar_N                            | fr_methoxy         |
| Chi0n                            | SlogP_VSA8 | fr_Ar_NH                           | fr_morpholine      |
| Chi1                             | SlogP_VSA9 | fr_Ar_OH                           | fr_nitrile         |
| Chi1n                            | TPSA     | fr_COO                             | fr_nitro           |
| Chi1v                            | Estate_VSA1 | fr_COO2                           | fr_nitro_arom      |
| Chi2n                            | Estate_VSA10 | fr_C_O                           | fr_nitro_arom_nonortho|
| Chi2v                            | Estate_VSA11 | fr_C_O_noCOO                      | fr_nitroso         |
| Chi3n                            | Estate_VSA2 | fr_C_S                            | fr_oxazole         |
| Chi3v                            | Estate_VSA3 | fr_HOCCN                           | fr_oxime           |
| Chi4n                            | Estate_VSA4 | fr_Imine                           | fr_para_hydroxylation|
| Chi4v                            | Estate_VSA5 | fr_NH0                             | fr_phenol          |
| HallKierAlpha                    | Estate_VSA6 | fr_NH1                             | fr_phenol_noOrthoHbond|
| Ipc                              | Estate_VSA7 | fr_NH2                             | fr_phos_acid       |
| Kappa1                           | Estate_VSA8 | fr_N_O                             | fr_phos_ester      |
| Kappa2                           | Estate_VSA9 | fr_Ndealkylation1                  | fr_piperdine       |
| Kappa3                           | VSA_EState1 | fr_Ndealkylation2                  | fr_piperazine      |
| LabuteASA                        | VSA_EState10 | fr_Nhprrrole                       | fr_priamide        |
| PEOE_VSA1                        | VSA_EState2 | fr_SH                              | fr_priulfonamid    |
| PEOE_VSA10                       | VSA_EState3 | fr_aldehyde                        | fr_pyridine        |
| PEOE_VSA11                       | VSA_EState4 | fr_alkyl_carbamate                 | fr_quatN           |
| PEoe_VSA12 | VSA_EState5 | fr_alkyl_halide | fr_sulfide |
| PEoe_VSA13 | VSA_EState6 | fr_allylic_oxid | fr_sulfonamid |
| PEoe_VSA14 | VSA_EState7 | fr_amide | fr_sulfone |
| PEoe_VSA2  | VSA_EState8 | fr_amidine | fr_term_acetylene |
| PEoe_VSA3  | VSA_EState9 | fr_aniline | fr_tetrazole |
| PEoe_VSA4  | FractionCSP3 | fr_aryl_methyl | fr_thiazole |
| PEoe_VSA5  | HeavyAtomCount | fr_azide | fr_thiocyan |
| PEoe_VSA6  | NHOHCount | fr_azo | fr_thiophene |
| PEoe_VSA7  | NOCount | fr_barbitur | fr_unbrch_alkane |
| PEoe_VSA8  | NumAliphaticCarbocycles | fr_benzene | fr_urea |

Table S1. List of function names for molecular descriptor calculation in RDKit (65).
**In Silico Generation of Dataset C**

For the generation of Dataset C, polyimides can be synthesized through the reaction between a dianhydride and a diamine with the elimination of water molecules, or through the reaction between a dianhydride and a diisocyanate with the elimination of carbon dioxide molecules, shown in **Fig. S1(a)**. All of the polyimides in Dataset C are generated via this mechanism from binary pairs of known chemicals in PubChem.

Double-stranded ladder polymers can be obtained through the binary reaction of a tetra-halogenated aromatic monomer and a tetra-hydroxy aromatic monomer, as shown in **Fig. S1(b)**. We obtain ~500 ladder polymers from different binary combinations of selected reacting monomers, which defines a chemical space of ladder polymers with sparse data points. To better explore this chemical space, we train a RNN model to learn its structure patterns and populate another 500 ladder polymers to supplement the chemical space. Together, the ladder polymers from monomer combinations and RNN generation comprise Dataset D.

![Fig. S1](image-url) **Fig. S1. In-silico reaction routes for the generation of polyimides (Dataset C) and ladder polymers (Dataset D).** (a) The polycondensation of diisocyanate/diamine and dianhydride to form a polyimide. Each generated polyimide features two functional imide groups. (b) The formula for ladder polymer formation from a tetra-halogenated aromatic monomer and a tetra-hydroxy aromatic monomer. Each ladder polymer repeating unit features four connection points along its double stranded backbone. Reacting groups are highlighted in yellow.
Fig. S2. Visualization of the MFF chemical space for the datasets explored in this work, using uniform manifold approximation and projection (UMAP), which captures local and global structure in the data. (a-d) correspond to Datasets A-D respectively, each compared to the PoLyInfo database, one of the largest polymer databases. Dataset A (training set) generally spans the feature space occupied by Datasets B, C, and D.
Fig. S3. Further visualization of imputed permeabilities in the training set, Dataset A, from the PoLyInfo and the MSA Databases. (a) O$_2$/N$_2$, (b) CO$_2$/CH$_4$, (c) CO$_2$/N$_2$ and (d) H$_2$/CO$_2$ Robeson plots after imputation using extremely randomized trees (ERT) and Bayesian linear regression (BLR), with permeabilities averaged across entries that correspond to the same polymer. The gas permeabilities of all six gases (He, H$_2$, O$_2$, N$_2$, CO$_2$, and CH$_4$) as visualized in lower-dimensional space using (e) principal components analysis (PCA).
Fig. S4. Summary of the performance of the random forest model trained on MFFs and gas permeabilities imputed using BLR. The predicted and actual permeabilities in Barrers for six gases are plotted for the train and test datasets.
Fig. S5. Summary of the performance of the ensemble of DNNs trained on MFFs and gas permeabilities imputed using BLR. The predicted and actual permeabilities in Barrers for six gases are plotted for the train and test datasets.
|                | He    | H₂    | O₂    | N₂    | CO₂   | CH₄   |
|----------------|-------|-------|-------|-------|-------|-------|
| RMSE train set | 0.296 | 0.327 | 0.366 | 0.405 | 0.387 | 0.453 |
| RMSE test set  | 0.259 | 0.309 | 0.340 | 0.410 | 0.383 | 0.495 |
| RMSE blind set | 0.366 | 0.388 | 0.473 | 0.549 | 0.439 | 0.566 |
| Mean single-model test R² | 0.69  | 0.68  | 0.71  | 0.68  | 0.70  | 0.70  |
| Std of test R² | 0.043 | 0.070 | 0.049 | 0.098 | 0.043 | 0.054 |
| Ensemble test R² | 0.91  | 0.90  | 0.92  | 0.91  | 0.90  | 0.88  |
| Mean variance of normalized test predictions | 0.071 | 0.086 | 0.072 | 0.101 | 0.073 | 0.074 |

Table S2. Summary of performance and uncertainty quantification for the ensemble of 16 DNNs trained on MFFs and permeabilities imputed using BLR. Performance on three different datasets is measured by the root mean squared error (RMSE) normalized by the standard deviation of permeabilities in the training set. The train and test sets are drawn from a shared distribution (from MSA and PoLyInfo). However, the blind set consists of 74 other high-performance polymers in the literature, which are assumed to be drawn from an unseen distribution. The train and test sets have similar normalized RMSEs, which suggests that the DNN model generalizes well to the distribution of polymers from MSA and PoLyInfo. The errors are about 25% higher when performing zero-shot learning on the blind set, which suggests that the DNN model still generalizes relatively well to completely unseen polymers. The last row refers to variance on predicted permeability values that have been normalized to a standard deviation of 1 and a mean of 0. The polymers used in the blind set can be found in the data repository associated with this work: [http://dx.doi.org/10.22002/D1.20048](http://dx.doi.org/10.22002/D1.20048).
| ID  | Descriptor Name       | Description                                                                 |
|-----|----------------------|-----------------------------------------------------------------------------|
| 101 | VSA_EState8          | Hybrid MOE-type descriptor using EState indices and VSA contributions for 6.45 < x < 7.00 |
| 107 | NumAliphaticCarbocycles | Number of carbocycles that contain at least one non-aromatic bond         |
| 109 | NumAliphaticRings     | Number of rings that contain at least one non-aromatic bond                  |
| 15  | FpDensityMorgan2      | Approximated density of the molecule                                        |
| 75  | SlogP_VSA4           | Sum of the approximate accessible VDW surface areas of atoms in the molecule with contributions between 0 and 0.1 to the molecule octanol/water partition coefficient calculation method proposed by Crippen |
| 79  | SlogP_VSA8           | Sum of the approximate accessible VDW surface areas of atoms in the molecule with contributions between 0.25 and 0.3 to the molecule octanol/water partition coefficient calculation method proposed by Crippen |
| 103 | FractionCSP3         | Fraction of C atoms that are sp3 hybridized                                 |
| 91  | EState_VSA8          | Hybrid MOE-type descriptor using EState indices and VSA contributions for 2.05 < x < 4.69 |
| 16  | FpDensityMorgan3      | Approximated density of the molecule                                        |
| 69  | SlogP_VSA1           | Sum of the approximate accessible VDW surface areas of atoms in the molecule with contributions less than -0.40 to the molecule octanol/water partition coefficient calculation method proposed by Crippen |
| 92  | EState_VSA9          | Hybrid MOE-type descriptor using EState indices and VSA contributions for 4.69 < x < 9.17 |
| 65  | SMR_VSA6             | MOE-type descriptor using surface area contributions where 2.75 < x < 3.05    |

Table S3. Definitions of the top molecular descriptors as identified using SHAP on the DNN model trained on descriptors with permeabilities imputed using BLR. All descriptors are calculated based on the graph structure and atom and bond types within the molecule. EState refers to the electrotopological state indices calculated as proposed by Hall and Kier (69). Large positive values indicate atoms of high electronegativity and/or terminal atoms, while small or negative EState values indicate atoms with only \( \sigma \) electrons, interior atoms, or atoms near electronegative atoms. VSA refers to the accessible van der Waals surface area.
Fig. S6. SHAP summary plots showing the impacts of the twenty most important chemical descriptors. We consider (a) He, (b) H$_2$, (c) O$_2$, (d) N$_2$, (e) CO$_2$, and (f) CH$_4$ gas permeabilities in the DNN ensemble model trained on descriptors with permeabilities imputed via BLR.
**Fig. S7. Correlations of top molecular descriptors.** The descriptors can be broken down into two main opposing correlation groups: [15, 16] are anti-correlated with [107, 109], whereas the other descriptors are relatively independent.
Fig. S8. SHAP summary plots showing the impacts of the twenty most important chemical substructures. We consider (a) He, (b) H₂, (c) O₂, (d) N₂, (e) CO₂, and (f) CH₄ gas permeabilities in the DNN ensemble model trained on MFFs with permeabilities imputed via BLR.
Fig. S9. Correlations of top molecular fingerprints. The main correlation group is [822, 1432, 1781].
Fig. S10. Visualization of predicted permeabilities for polymers in Datasets B, C, and D, based on the RF trained on MFFs with BLR-imputed permeabilities. The training dataset (Dataset A) is overlaid on the predicted permeabilities. The data is visualized for (a) O$_2$/N$_2$, (b) CO$_2$/CH$_4$, (c) CO$_2$/N$_2$, and (d) H$_2$/CO$_2$ separations. Units of permeability are Barrers.
| Name       | Polymer | Component A                  | Component B                  |
|------------|---------|------------------------------|------------------------------|
| P-DNN-C1   | ![Diagram](P-DNN-C1.png) | ![Structure](P-DNN-C1_STRUCTURE.png) | ![Structure](P-DNN-C1_STRUCTURE_B.png) |
| P-DNN-C2   | ![Diagram](P-DNN-C2.png) | ![Structure](P-DNN-C2_STRUCTURE.png) | ![Structure](P-DNN-C2_STRUCTURE_B.png) |
| P-DNN-C3   | ![Diagram](P-DNN-C3.png) | ![Structure](P-DNN-C3_STRUCTURE.png) | ![Structure](P-DNN-C3_STRUCTURE_B.png) |
P-DNN-C4

\*N(C(C)=O)C(=O)CS(=O)(=O)CC(=O)N(C(C)=O)C(C)C1(*)CCCCCCCCC

P-DNN-D1

\*Oc1cc2c(cc1O*)C1(c3cc4c(cc3-
c3cc5c(cc31)C(C)(C)CCC5(C)C)Oc1c(c(C(=
O)c3cc(F)c(F)c(F)c3)c(*)c(*)c1C(=O)c1cc(F)
ce(F)e1)O4)e1cc3c(e1-2)C(C)(C)CC3(C)C

P-DNN-D2

\*Oc1cc2c(cc1O*)C1(c3cc4c(cc3-
c3cc5c(cc31)C(C)(C)CCC5(C)C)Oc1c(c(C(=
O)c3cc(F)c(F)c(F)c3)c(*)c(*)c1C(=O)c1cc(F)
ce(F)e1)O4)e1cc3c(e1-2)C(C)(C)CC3(C)C

P-RF-C

\*e1eC)cc2cccccc2c1-
c1e(N2C(=O)c3cc4e(cc3C2=O)C2(C)CC4(C)
c3cc4e(cc32)C(=O)N(*)C4=O)e(c(C)cc2cccccc1

SA: 5.65
AS: -8.36, -6.76

SA: 3.06
AS: -1.29, -1.04

SA: 1.99
AS: -7.97, -8.74

SA: 2.55
AS: -6.5, -9.27

SA: 4.00
AS: -4.81, -6.66

SA: 2.29
AS: -5.81, -5.61

SA: 1.99
AS: -7.97, -8.74
Table S4. Selected top polymer candidates and the individual components required for their syntheses. Polymers are labeled by their SMILES strings, associated synthetic accessibility (SA) scores (53), and aqueous solubilities (AS). SAs are calculated using an algorithm that accounts for fragment contributions and a complexity penalty, producing a rating between 1 (easy to synthesize) and 10 (very difficult to synthesize); this method has been validated by ratings given by expert medicinal chemists. We use AqSolPred (70) and SolTranNet (71) to estimate the AS of components A and B. These two methods are essentially supervised machine learning models trained on drug-like molecules to predict solubility values (LogS).
Fig. S11. Distributions of the synthetic accessibility (SA) scores. (a) diamines/diisocynates and (b) dianhydrides required for the formation of the promising polyimides summarized in Table 3 of the main text.
Solubility estimations for polymers

We further utilize the Polymer Genome (72) to predict the solubilities of polymer candidates in organic solvents based on the Hildebrand solubility parameter. Essentially, the Polymer Genome is a machine learning model that predicts various properties of polymers, trained on properties from first-principles computations and experimental measurements. Several solvent options are found to be suitable for each selected polymer candidate, as summarized in Table S5-12. These estimations can provide guidelines for experimentalists to select a proper solvent for fabrication.

| Solvent                                | Solubility parameter | Non-Solvent                               | Solubility parameter |
|----------------------------------------|----------------------|-------------------------------------------|----------------------|
| Butyric Acid                           | 18.7 MPa$^{1/2}$     | 3-Methoxypropyl Amine/Acetic Acid         | 36.8 MPa$^{1/2}$     |
| Methyl Acetate                         | 18.7 MPa$^{1/2}$     | 2-Amino-2-Methyl-1-Propanol/Acetic Acid   | 36.7 MPa$^{1/2}$     |
| Di-Isobutyl Carbinol                   | 18.7 MPa$^{1/2}$     | Formamide                                 | 36.7 MPa$^{1/2}$     |
| Furan                                  | 18.7 MPa$^{1/2}$     | Dimethyl Ethanolamine/Formic Acid         | 35.6 MPa$^{1/2}$     |
| Cyclohexylamine                        | 18.7 MPa$^{1/2}$     | Dimethyl Ethanolamine/p-Toluene Sulfonic Acid | 35.6 MPa$^{1/2}$  |
| Ethylene Glycol Monoisobutyl Ether     | 18.6 MPa$^{1/2}$     | Dimethyl Ethanolamine/Thioglycolic Acid   | 35.6 MPa$^{1/2}$     |
| Hexafluoro hexanol                     | 18.6 MPa$^{1/2}$     | Ethylene Glycol                           | 33 MPa$^{1/2}$       |
| Chloroform                             | 19 MPa$^{1/2}$       | Dimethyl Ethanolamine/Acetic Acid         | 32.7 MPa$^{1/2}$     |
| Diethylene Glycol Methyl t-Butyl Ether | 19 MPa$^{1/2}$       | Triethanolamine/Acetic Acid               | 32.4 MPa$^{1/2}$     |
| Mesityl Oxide                          | 18.5 MPa$^{1/2}$     | Ethanolamine/Acetic Acid                  | 32.4 MPa$^{1/2}$     |

Table S5. Solubility summary for P-DNN-C1.
| Solvent                        | Solubility parameter | Non-Solvent                                      | Solubility parameter |
|-------------------------------|----------------------|--------------------------------------------------|----------------------|
| Methyl Methacrylate           | 17.9 MPa$^{1/2}$     | 3-Methoxypropyl Amine/Acetic Acid                | 36.8 MPa$^{1/2}$     |
| Ethyl Acrylate                | 17.9 MPa$^{1/2}$     | 2-Amino-2-Methyl-1-Propanol/Acetic Acid          | 36.7 MPa$^{1/2}$     |
| p-Xylene                      | 17.9 MPa$^{1/2}$     | Formamide                                        | 36.7 MPa$^{1/2}$     |
| Ethylene Glycol Methyl t-Butyl Ether* | 18.1 MPa$^{1/2}$ | Dimethyl Ethanolamine/p-Toluene Sulfonic Acid | 35.6 MPa$^{1/2}$     |
| Carbon Tetrachloride          | 18.1 MPa$^{1/2}$     | Dimethyl Ethanolamine/Formic Acid                | 35.6 MPa$^{1/2}$     |
| Ethyl Benzene                 | 17.9 MPa$^{1/2}$     | Dimethyl Ethanolamine/Thioglycolic Acid          | 35.6 MPa$^{1/2}$     |
| Carbon Dioxide                | 17.9 MPa$^{1/2}$     | Ethylene Glycol                                  | 33 MPa$^{1/2}$       |
| Ethyl Acetate                 | 18.2 MPa$^{1/2}$     | Dimethyl Ethanolamine/Acetic Acid                | 32.7 MPa$^{1/2}$     |
| Triethylamine                 | 17.8 MPa$^{1/2}$     | Ethanolamine/Acetic Acid                         | 32.4 MPa$^{1/2}$     |
| Carbon tetrachloride          | 17.8 MPa$^{1/2}$     | Morpholine/Acetic Acid                           | 32.4 MPa$^{1/2}$     |

Table S6. Solubility summary for P-DNN-C2.
Table S7. Solubility summary for P-DNN-C3.

| Solvent                          | Solubility parameter | Non-Solvent                                    | Solubility parameter |
|---------------------------------|----------------------|------------------------------------------------|----------------------|
| Cyclohexanol                    | 22.4 MPa$^{1/2}$    | 3-Methoxypropyl Amine/Acetic Acid              | 36.8 MPa$^{1/2}$     |
| Aniline                         | 22.5 MPa$^{1/2}$    | 2-Amino-2-Methyl-1-Propanol/Acetic Acid        | 36.7 MPa$^{1/2}$     |
| Methacrylonitrile               | 22.5 MPa$^{1/2}$    | Formamide                                      | 36.7 MPa$^{1/2}$     |
| Acetic anhydride                | 22.3 MPa$^{1/2}$    | Dimethyl Ethanolamine/Formic Acid              | 35.6 MPa$^{1/2}$     |
| Benzyl Butyl Phthalate          | 22.3 MPa$^{1/2}$    | Dimethyl Ethanolamine/p-Toluene Sulfonic Acid  | 35.6 MPa$^{1/2}$     |
| Triethylphosphate               | 22.2 MPa$^{1/2}$    | Dimethyl Ethanolamine/Thioglycolic Acid        | 35.6 MPa$^{1/2}$     |
| Diethylene Glycol Monoethyl Ether | 22.2 MPa$^{1/2}$ | Ethylene Glycol                                | 33 MPa$^{1/2}$       |
| 2-Butanol                       | 22.2 MPa$^{1/2}$    | Dimethyl Ethanolamine/Acetic Acid              | 32.7 MPa$^{1/2}$     |
| Isobutyl Alcohol                | 22.7 MPa$^{1/2}$    | Morpholine/Acetic Acid                         | 32.4 MPa$^{1/2}$     |
| Nitrobenzene                    | 22.2 MPa$^{1/2}$    | Ethanolamine/Acetic Acid                       | 32.4 MPa$^{1/2}$     |
Table S8. Solubility summary for P-DNN-C4.

| Solvent                          | Solubility parameter | Non-Solvent                               | Solubility parameter |
|----------------------------------|----------------------|-------------------------------------------|----------------------|
| Ethylene Dichloride              | 20.8 MPa$^{1/2}$     | 3-Methoxypropyl Amine/Acetic Acid         | 36.8 MPa$^{1/2}$     |
| Ethylene Glycol Monobutyl Ether  | 20.8 MPa$^{1/2}$     | 2-Amino-2-Methyl-1-Propanol/Acetic Acid   | 36.7 MPa$^{1/2}$     |
| Diacetone Alcohol                | 20.8 MPa$^{1/2}$     | Formamide                                 | 36.7 MPa$^{1/2}$     |
| 2-Nitropropane                   | 20.6 MPa$^{1/2}$     | Dimethyl Ethanolamine/p-Toluene Sulfonic Acid | 35.6 MPa$^{1/2}$     |
| 1-Bromonaphthalene               | 20.9 MPa$^{1/2}$     | Dimethyl Ethanolamine/Formic Acid         | 35.6 MPa$^{1/2}$     |
| Carbon Disulfid                  | 20.5 MPa$^{1/2}$     | Dimethyl Ethanolamine/Thioglycolic Acid   | 35.6 MPa$^{1/2}$     |
| 1,4-Dioxane                      | 20.5 MPa$^{1/2}$     | Ethylene Glycol                           | 33 MPa$^{1/2}$       |
| o-Dichlorobenzene                | 20.5 MPa$^{1/2}$     | Dimethyl Ethanolamine/Acetic Acid         | 32.7 MPa$^{1/2}$     |
| Propylene Glycol Monomethyl Ether| 20.4 MPa$^{1/2}$     | Triethanolamine/Acetic Acid               | 32.4 MPa$^{1/2}$     |
| Diethylene Glycol Monobutyl Ether| 20.4 MPa$^{1/2}$    | Morpholine/Acetic Acid                    | 32.4 MPa$^{1/2}$     |
Table S9. Solubility summary for P-DNN-D1.

| Solvent                        | Solubility parameter | Non-Solvent                          | Solubility parameter |
|--------------------------------|----------------------|--------------------------------------|----------------------|
| 2-Ethyl-hexanol                | 20.1 MPa$^{1/2}$     | 3-Methoxypropyl Amine/Acetic Acid    | 36.8 MPa$^{1/2}$     |
| Methyl Isobutyl Carbinol       | 20 MPa$^{1/2}$       | 2-Amino-2-Methyl-1-Propanol/Acetic Acid | 36.7 MPa$^{1/2}$     |
| Dibutyl Phthalate              | 20.2 MPa$^{1/2}$     | Formamide                            | 36.7 MPa$^{1/2}$     |
| Dipropylene Glycol Methyl Ether| 20 MPa$^{1/2}$       | Dimethyl Ethanolamine/p-Toluene Sulfonic Acid | 35.6 MPa$^{1/2}$     |
| Methylene Dichloride           | 20.2 MPa$^{1/2}$     | Dimethyl Ethanolamine/Formic Acid    | 35.6 MPa$^{1/2}$     |
| Acetone                        | 19.9 MPa$^{1/2}$     | Dimethyl Ethanolamine/Thioglycolic Acid | 35.6 MPa$^{1/2}$     |
| Isophorone                     | 19.9 MPa$^{1/2}$     | Ethylene Glycol                      | 33 MPa$^{1/2}$       |
| Tetrahydronaphthalene          | 19.9 MPa$^{1/2}$     | Dimethyl Ethanolamine/Acetic Acid    | 32.7 MPa$^{1/2}$     |
| Butyl Lactate                  | 19.9 MPa$^{1/2}$     | Ethanolamine/Acetic Acid             | 32.4 MPa$^{1/2}$     |
| Butyronitrile                  | 20.3 MPa$^{1/2}$     | Triethanolamine/Acetic Acid          | 32.4 MPa$^{1/2}$     |

Predicted Hildebrand solubility parameter: 20.1 MPa$^{1/2}$
| Solvent                             | Solubility parameter | Non-Solvent                          | Solubility parameter |
|------------------------------------|----------------------|--------------------------------------|----------------------|
| 2,4-Pentanediode                   | 19.8 MPa\(^{1/2}\)   | 3-Methoxypropyl Amine/Acetic Acid    | 36.8 MPa\(^{1/2}\)   |
| Butyl Lactate                      | 19.9 MPa\(^{1/2}\)   | 2-Amino-2-Methyl-1-Propanol/Acetic Acid | 36.7 MPa\(^{1/2}\)   |
| Tetrahydronaphthalene              | 19.9 MPa\(^{1/2}\)   | Formamide                            | 36.7 MPa\(^{1/2}\)   |
| Ethylene Glycol Mono-t-Butyl Ether | 19.7 MPa\(^{1/2}\)   | Dimethyl Ethanolamine/Formic Acid    | 35.6 MPa\(^{1/2}\)   |
| Acetone                            | 19.9 MPa\(^{1/2}\)   | Dimethyl Ethanolamine/Thioglycolic Acid | 35.6 MPa\(^{1/2}\)   |
| Isophorone                         | 19.9 MPa\(^{1/2}\)   | Dimethyl Ethanolamine/p-Toluene Sulfonic Acid | 35.6 MPa\(^{1/2}\)   |
| Dipropylene Glycol Methyl Ether    | 20 MPa\(^{1/2}\)     | Ethylene Glycol                      | 33 MPa\(^{1/2}\)     |
| Ethylene Glycol Monoethyl Ether Acetate | 19.7 MPa\(^{1/2}\)   | Dimethyl Ethanolamine/Acetic Acid    | 32.7 MPa\(^{1/2}\)   |
| Butoxy Ethoxy Propanol             | 19.7 MPa\(^{1/2}\)   | Triethanolamine/Acetic Acid          | 32.4 MPa\(^{1/2}\)   |
| Methyl Isobutyl Carbinol           | 20 MPa\(^{1/2}\)     | Morpholine/Acetic Acid               | 32.4 MPa\(^{1/2}\)   |

Table S10. Solubility summary for P-DNN-D2.
| Solvent                        | Solubility parameter | Non-Solvent                                      | Solubility parameter |
|-------------------------------|----------------------|-------------------------------------------------|----------------------|
| Dimethyl Phthalate            | 22.1 MPa$^{1/2}$     | 3-Methoxypropyl Amine/Acetic Acid               | 36.8 MPa$^{1/2}$     |
| Tetrahydrofurfuryl Alcohol    | 22.1 MPa$^{1/2}$     | 2-Amino-2-Methyl-1-Propanol/Acetic Acid         | 36.7 MPa$^{1/2}$     |
| Nitrobenzene                  | 22.2 MPa$^{1/2}$     | Formamide                                       | 36.7 MPa$^{1/2}$     |
| Diethylene Glycol Monoethyl Ether | 22 MPa$^{1/2}$  | Dimethyl Ethanolamine/p-Toluene Sulfonic Acid   | 35.6 MPa$^{1/2}$     |
| 2-Butanol                     | 22.2 MPa$^{1/2}$     | Dimethyl Ethanolamine/Formic Acid               | 35.6 MPa$^{1/2}$     |
| Diethylene Glycol Monoethyl Ether | 22.2 MPa$^{1/2}$ | Dimethyl Ethanolamine/Thioglycolic Acid         | 35.6 MPa$^{1/2}$     |
| Triethylphosphate             | 22.2 MPa$^{1/2}$     | Ethylene Glycol                                 | 33 MPa$^{1/2}$       |
| 1-Pentanol                    | 21.9 MPa$^{1/2}$     | Dimethyl Ethanolamine/Acetic Acid               | 32.7 MPa$^{1/2}$     |
| Benzyl Butyl Phthalate        | 22.3 MPa$^{1/2}$     | Morpholine/Acetic Acid                          | 32.4 MPa$^{1/2}$     |
| Acetic anhydride              | 22.3 MPa$^{1/2}$     | Ethanolamine/Acetic Acid                        | 32.4 MPa$^{1/2}$     |

Table S11. Solubility summary for P-RF-C.
Predicted Hildebrand solubility parameter: 23.4 MPa$^{1/2}$

| Solvent                        | Solubility parameter | Non-Solvent                        | Solubility parameter |
|--------------------------------|----------------------|------------------------------------|----------------------|
| Hexamethylphosphoramide       | 23.3 MPa$^{1/2}$     | 3-Methoxypropyl Amine/Acetic Acid | 36.8 MPa$^{1/2}$     |
| Ethylene Glycol Monoethyl Ether | 23.5 MPa$^{1/2}$    | 2-Amino-2-Methyl-1-Propanol/Acetic Acid | 36.7 MPa$^{1/2}$ |
| 3-Chloro-1-Propanol           | 23.6 MPa$^{1/2}$     | Formamide                          | 36.7 MPa$^{1/2}$     |
| 1-Butanol                     | 23.2 MPa$^{1/2}$     | Dimethyl Ethanolamine/Formic Acid  | 35.6 MPa$^{1/2}$     |
| 2-Propanol                    | 23.6 MPa$^{1/2}$     | Dimethyl Ethanolamine/Thioglycolic Acid | 35.6 MPa$^{1/2}$ |
| Tricresyl Phosphate           | 23.1 MPa$^{1/2}$     | Dimethyl Ethanolamine/p-Toluene Sulfonic Acid | 35.6 MPa$^{1/2}$ |
| Hexafluoroisopropanol         | 23.1 MPa$^{1/2}$     | Ethylene Glycol                    | 33 MPa$^{1/2}$       |
| Benzyl Alcohol                | 23.8 MPa$^{1/2}$     | Dimethyl Ethanolamine/Acetic Acid  | 32.7 MPa$^{1/2}$     |
| N-Methyl-2-Pyrrolidone        | 23 MPa$^{1/2}$       | Morpholine/Acetic Acid             | 32.4 MPa$^{1/2}$     |
| N,N-Dimethyl Acetamide        | 22.8 MPa$^{1/2}$     | Ethanolamine/Acetic Acid           | 32.4 MPa$^{1/2}$     |

**Table S12. Solubility summary for P-RF-D.**
| Polymer          | # Gas mols: | Gas Number/ Diffusivity [10^-8 cm^2/s] | Solubility [cm^3 (STP)/cm^2 bar] |
|------------------|-------------|---------------------------------------|----------------------------------|
|                  |             | 10        | 20        | 30        | 40        | 50        |                     |
| P-DNN-C1         | CH4         | 44.6067   | 60.5667   | 59.4783   | 58.085    | 99.6317   | 21.473              |
|                  | CO2         | 79.935    | 190.8     | 181.9     | 190.9933  | 261.8417  | 69.456              |
|                  | N2          | 53.1483   | 79.9667   | 82.9083   | 82.6283   | 124.0167  | 13.161              |
|                  | O2          | 283.4     | 439.5333  | 454.55    | 445.4833  | 658.0833  | 10.67               |
|                  | H2          | 5085.1768 | 8706.2073 | 8612.0009 | 8318.7770 | 9630.0227 | 0.827               |
| P-DNN-C2         | CH4         | 1066.467  | 2940.167  | 3008.667  | 3069.833  | 4436.917  | 33.832              |
|                  | CO2         | 2457.167  | 9288.333  | 8900      | 8613.333  | 11448.17  | 75.4                |
|                  | N2          | 1186.283  | 2080      | 2329.5    | 2203.667  | 2603.667  | 13.843              |
|                  | O2          | 3436.667  | 5883.833  | 5820.167  | 5622      | 6508.167  | 9.081               |
|                  | H2          | 18510.480 | 25402.7283| 25005.3350| 25673.7933| 31672.1650| 1.359               |
| P-DNN-C3         | CH4         | 607.9195  | 934.1333  | 974.1057  | 900.75    | 1510.1495 | 57.225              |
|                  | CO2         | 16001.2953| 27005.769 | 28008.8731| 27506.6922| 45003.2976| 1.419               |
|                  | N2          | 365.1723  | 632.5788  | 613.1065  | 653.1013  | 904.6467  | 10.017              |
|                  | O2          | 4008.6947 | 6204.4736 | 6608.0365 | 6406.684  | 9503.2363 | 6.088               |
|                  | H2          | 450.3511  | 709.6833  | 632.9442  | 669.0403  | 1258.4709 | 33.926              |
| P-DNN-C4         | CH4         | 2710.163  | 4709.488  | 4501.601  | 4654.203  | 8001.208  | 34.534              |
|                  | CO2         | 30007.729 | 47002.976 | 45508.835 | 47506.752 | 70008.792 | 1.983               |
|                  | N2          | 347.1227  | 852.2008  | 822.1247  | 862.3635  | 1509.3613 | 8.135               |
|                  | O2          | 7008.975  | 10503.245 | 10005.407 | 9504.424  | 18007.930 | 5.424               |
|                  | H2          | 457.240   | 1359.953  | 1282.746  | 1486.643  | 1808.290  | 19.443              |
| P-DNN-D1         | CH4         | 290.7     | 250.95    | 252.375   | 270.325   | 399.325   | 5.700               |
|                  | CO2         | 125.1567  | 163.38    | 154.735   | 156.5817  | 163.5967  | 75.101              |
|                  | N2          | 202.4     | 421.8833  | 414.8667  | 432.25    | 791.7167  | 3.68                |
|                  | O2          | 619.4833  | 1899      | 1863.5    | 1949      | 2496.5    | 1.518               |
|                  | H2          | 9999.3383 | 15892.3700| 15706.9133| 16479.6533| 19980.1867| 0.284               |
| P-DNN-D2         | CH4         | 303.4167  | 480.8667  | 428.8333  | 537       | 549.9     | 5.922               |
|                  | CO2         | 652.225   | 1122.87   | 1081.99   | 1012.725  | 3161.725  | 17.876              |
|                  | N2          | 482.5     | 1052.2    | 1023.517  | 1046.45   | 1834.833  | 3.581               |
|                  | O2          | 1226.667  | 2994.667  | 2666.667  | 2752.167  | 3828.833  | 2.175               |
|                  | H2          | 14993.9500| 21997.9900| 20999.5850| 23004.1967| 27019.6850| 0.389               |
| P-RF-C           | CH4         | 12.9317   | 12.33     | 11.9883   | 13.2017   | 35.76     | 16.050              |
|                  | CO2         | 30.8567   | 30.74     | 32.5      | 34.9843   | 101.59    | 70.120              |
|                  | N2          | 5.9833    | 12.7667   | 10.8333   | 12.0333   | 19.4883   | 10.941              |
|                  | O2          | 36.65     | 64.86     | 73.9517   | 70.12     | 78.1167   | 6.575               |
|                  | H2          | 1001.2217 | 1902.8250 | 2058.2783 | 2002.8350 | 2890.8383 | 0.716               |
| P-RF-D           | CH4         | 1.8483    | 2.3267    | 3.02      | 2.745     | 6.98      | 53.742              |
|                  | CO2         | 8.9933    | 9.8783    | 8.4167    | 8.4933    | 29.4      | 218.717             |
|                  | N2          | 1.9       | 3.0042    | 3.9547    | 3.3403    | 6.6538    | 29.255              |
|                  | O2          | 15.8738   | 33.63     | 33.0467   | 37.5233   | 64.1417   | 13.437              |
|                  | H2          | 644.5450  | 1201.9217 | 1185.6833 | 1244.1850 | 1634.6150 | 1.203               |

Table S13. Intermediate values in the calculation of permeability as diffusivity times solubility for selected top polymer candidates, via MD simulations with the COMPASS forcefield (73). The diffusivity was calculated as the average of the shaded values, which refer to simulations with 20, 30, and 40 gas molecules.
| PIM-1 | Gas | \([\text{cm}^3 \text{ (STP)/cm}^3 \text{ (polymer bar)}]\) | \([10^{-8} \text{ cm}^2/\text{s}]\) | Permeability |
|-------|-----|------------------|-----------------|--------------|
|       |     | \(S_{\text{sim}}\) | \(D_{\text{sim}}\) | \(S_{\text{exp}}\) (46) | \(D_{\text{exp}}\) (46) | \(P_{\text{exp}}\) (46) |
|       | Our work | Ref. (23) | Our work | Ref. (23) | Our work | |
| CH\(_4\) | 15.1±1.91 | 14.2±3.1 | 13.70 | 8.05±0.23 | 112±27 | 6.80 | 155.147±10.71 | 121.108 |
| CO\(_2\) | 49.93±3.41 | 50.7±9 | 66.90 | 28.45±4.19 | 151±47 | 26.00 | 2150.966±200.718 | 2261.22 |
| H\(_2\) | 0.42±0.06 | 0.46±0.03 | 0.58 | 1826.33±43.33 | 6630±215 | 1700.00 | 983.104±121.417 | 1281.8 |
| N\(_2\) | 7.31±0.89 | - | - | 31.60±3.51 | - | - | 289.318±26.73 | - |
| O\(_2\) | 3.93±0.68 | 4.11±0.58 | 3.50 | 82.31±5.82 | 452±81 | 81.00 | 400.507±56.293 | 368.55 |

Table S14. Benchmark MD simulations for the PIM-1 polymer membrane. Simulated solubility coefficients \([\text{cm}^3 \text{ (STP)/cm}^3 \text{ (polymer) bar]}\), diffusion coefficients \([10^{-8} \text{ cm}^2/\text{s}]\) at 300K, and permeabilities \([\text{Barrer}]\) are given. The values from our study show good agreement with simulation and experimental values in the literature.
Fig. S12. Ten relevant polymers used for further benchmarking of MD simulations, using the method outlined in this study. (a) Chemical components and (b) atomistic models of the respective in-silico membranes for these polymers.
| No | Polymer | Gas type | Simulation value | Experiment value | Ref. |
|----|---------|----------|------------------|------------------|-----|
|    |         |          | Unit: Barrer     | Unit: Barrer     |     |
| 1  | 6FDA+3MPDA | H₂       | 396.7708 ± 11.7018 | 344              | (13) |
|    |          | CO₂      | 25.9387 ± 1.9684  | 15               |     |
|    |          | CH₄      | 0.72062 ± 0.1422  | -                |     |
|    |          | N₂       | 8.73845 ± 0.81306 | 5.2              |     |
|    |          | O₂       | 59.49297 ± 5.4821 | 62.5             |     |
| 2  | BTDA+4MPDA | H₂       | 129.5787 ± 1.8292 | 108              | (13) |
|    |          | CO₂      | 0.73685 ± 0.0305  | 0.428            |     |
|    |          | CH₄      | 3.30228 ± 0.61435 | -                |     |
|    |          | N₂       | 21.6524 ± 7.36388 | -                |     |
|    |          | O₂       | 13.2569 ± 1.2996  | 11.8             |     |
| 3  | HQDPA+3MPDA | H₂       | 48.8671 ± 2.8231  | 40               | (13) |
|    |          | CO₂      | 27.9553 ± 2.0265  | 18.2             |     |
|    |          | CH₄      | 2.35364 ± 0.2153  | -                |     |
|    |          | N₂       | 1.24346 ± 0.27718 | 0.9              |     |
|    |          | O₂       | 5.8980 ± 0.3891   | 4.42             |     |
| 4  | 6FDA+DABA | H₂       | 32.8373 ± 1.7064  | 25               | (74) |
|    |          | CO₂      | 5.96324 ± 0.8199  | 3.4              |     |
|    |          | CH₄      | 0.07622 ± 0.0189  | 0.054            |     |
|    |          | N₂       | 0.20209 ± 0.05512 | 0.126            |     |
|    |          | O₂       | 1.29972 ± 0.07917 | 1.01             |     |
| 5  | DPPD+IMM | H₂       | 1210.6574 ± 37.7677 | -         | (13) |
|    |          | CO₂      | 416.8465 ± 16.9437 | 392       |     |
|    |          | CH₄      | 30.8792 ± 1.5332  | 24               |     |
|    |          | N₂       | 24.4127 ± 0.9755  | 19               |     |
|    |          | O₂       | 85.08897 ± 4.6695 | 75               |     |
| 6  | 6FDA+SA55 | H₂       | 591.9396 ± 48.1532 | -         | (13) |
|    |          | CO₂      | 166.0332 ± 10.2192 | 149.72   |     |
|    |          | CH₄      | 9.13827 ± 0.76672 | 6.11             |     |
|    |          | N₂       | 6.16553 ± 0.41005 | 4.87             |     |
|    |          | O₂       | 7.74326 ± 0.8687  | 5.79             |     |
| 7  | Ladder A | H₂       | 4972.909 ± 168.8742 | 4520     | (75) |
|    |          | CO₂      | 10124.307 ± 221.659 | 8890    |     |
|    |          | CH₄      | 797.9956 ± 96.9325 | 622               |     |
|    |          | N₂       | 497.1791 ± 44.0883 | 441.875           |     |
|    |          | O₂       | 1825.66322 ± 156.654 | 1624      |     |
| 8  | Ladder B | H₂       | 2861.160 ± 181.16214 | 2520    | (76) |
|    |          | CO₂      | 2964.663 ± 493.1082 | 2741    |     |
|    |          | CH₄      | 3374.360 ± 293.1177 | 337               |     |
|    |          | N₂       | 246.25352 ± 19.04919 | 220       |     |
|    |          | O₂       | 895.8939 ± 14.13505 | 813               |     |
|   | Ladder C | H₂      | 591.9765 ± 33.67194 | 530 |
|   |         | CO₂     | 1337.6652 ± 126.0247 | 1100 |
|   |         | CH₄     | 90.3421 ± 10.0136    | 77  |
|   |         | N₂      | 72.8467 ± 3.0673     | 47  |
|   |         | O₂      | 163.8977 ± 23.8086   | 150 |
|   | Ladder D | H₂      | 721.01859 ± 15.50565 | 859 |
|   |         | CO₂     | 2133.1443 ± 200.727  | 1747 |
|   |         | CH₄     | 189.53161 ± 35.42094 | 148 |
|   |         | N₂      | 82.11123 ± 6.06716   | 76  |
|   |         | O₂      | 361.03137 ± 54.31964 | 273 |

Table S15. Comparison of MD-simulated permeabilities with experimental values for the ten selected polymers [Barrer: \(10^{-10} \text{cm}^3\text{STP cm}^2\text{cm}^{-3}\text{s}^{-1}\text{cmHg}\) at 300 K. Overall, our MD simulations agree well with experimentally measured gas permeabilities from the literature.
Table S16. Comparison of gas permeabilities calculated from MD simulations with permeabilities predicted from ML models, with error and uncertainty quantification, of selected top polymer candidates. Units of permeability are Barrers.

| Polymer | Gas | MD-Simulated Permeability (PMD) | Standard Deviation of PMD | ML-Predicted Permeability (PML) | log(PML) | Standard Deviation of log(PML) from Bootstrapping | %Difference (PML - PMD)/PMD |
|---------|-----|---------------------------------|---------------------------|---------------------------------|----------|-------------------------------------------------|-----------------------------|
| P-DNN-C1 | CH4 | 1657.512                         | 84.81211                  | 1653.12891                      | 3.218306 | 2.019868                                        | -0.26%                      |
|         | CO2 | 16965.81                         | 602.6442                  | 17662.1029                      | 4.247042 | 1.803747                                        | 4.10%                       |
|         | N2  | 1400.094                         | 46.43858                  | 1367.46                         | 3.135914 | 2.210223                                        | -2.33%                      |
|         | O2  | 6193.831                         | 158.9193                  | 6251.10322                      | 3.795956 | 1.777441                                        | 0.92%                       |
|         | H2  | 4727.856                         | 123.3328                  | 4685.6193                       | 3.670767 | 1.803201                                        | -0.89%                      |
| P-DNN-C2 | CH4 | 52486.33                         | 1830.314                  | 59462.9225                      | 4.774246 | 2.756414                                        | 13.29%                      |
|         | CO2 | 155897.4                         | 5214.004                  | 33963.203                       | 5.531011 | 2.413820                                        | 117.86%                     |
|         | N2  | 38486.6                          | 3429.304                  | 49003.579                       | 4.69022  | 2.302411                                        | 27.33%                      |
|         | O2  | 100832.1                         | 8783.601                  | 125618.927                      | 5.09905  | 2.180247                                        | 24.58%                      |
|         | H2  | 44794.08                         | 1108.444                  | 45040.167                       | 4.6536   | 1.950639                                        | 0.55%                       |
| P-DNN-C3 | CH4 | 69655.9                          | 2230.851                  | 66988.469                       | 4.826    | 3.704102                                        | -3.83%                      |
|         | CO2 | 50742.37                         | 755.4335                  | 51789.4144                      | 4.714241 | 2.526167                                        | 2.06%                       |
|         | N2  | 8242.063                         | 212.6468                  | 7448.64026                      | 3.872077 | 3.014252                                        | -9.63%                      |
|         | O2  | 50702.8                          | 1303.929                  | 53790.1803                      | 4.730703 | 3.087153                                        | 6.09%                       |
|         | H2  | 29574.06                         | 1382.519                  | 32884.0273                      | 4.516985 | 2.570558                                        | 11.19%                      |
| P-DNN-C4 | CH4 | 207490.4                         | 3946.849                  | 215341.632                      | 5.333128 | 4.503192                                        | 3.78%                       |
|         | CO2 | 120318                           | 2187.07                   | 126352.547                      | 5.101584 | 3.771231                                        | 5.02%                       |
|         | N2  | 8942.252                         | 180.6803                  | 9603.36782                      | 3.982514 | 4.410758                                        | 7.42%                       |
|         | O2  | 70542.73                         | 2875.251                  | 80745.9968                      | 4.907121 | 3.758369                                        | 14.46%                      |
|         | H2  | 26738.73                         | 1632.816                  | 22549.9197                      | 4.353145 | 3.595450                                        | -15.67%                     |
| P-DNN-D1 | CH4 | 1911.04                          | 70.41345                  | 1790.3582                       | 3.23283  | 0.799550                                        | -10.55%                     |
|         | CO2 | 15448.42                         | 371.7476                  | 14724.1268                      | 4.168029 | 0.800241                                        | -4.69%                      |
|         | N2  | 2023.817                         | 60.26185                  | 2117.81715                      | 3.325888 | 0.920442                                        | 4.64%                       |
|         | O2  | 3757.197                         | 83.67529                  | 3673.05404                      | 3.565263 | 0.824867                                        | -2.19%                      |
|         | H2  | 5915.074                         | 158.5987                  | 5925.44539                      | 3.772721 | 0.567156                                        | 0.18%                       |
| P-DNN-D2 | CH4 | 3712.361                         | 354.0196                  | 4066.00857                      | 3.609168 | 0.862484                                        | 9.53%                       |
|         | CO2 | 24924.16                         | 1507.761                  | 26067.0875                      | 4.416092 | 0.792417                                        | 4.59%                       |
|         | N2  | 4844.493                         | 100.834                   | 4733.66914                      | 3.67519  | 0.924626                                        | -2.29%                      |
|         | O2  | 7929.709                         | 402.0509                  | 7832.41445                      | 3.893895 | 0.754249                                        | -1.23%                      |
|         | H2  | 11111.7                          | 461.5145                  | 11813.6745                      | 4.072385 | 0.628416                                        | 6.32%                       |
| P-RF-C  | CH4 | 260.9494                         | 14.06993                  | 164.128095                      | 2.215182 | n/a                                             | -37.10%                     |
|         | CO2 | 2983.488                         | 168.798                   | 2713.03121                      | 3.433454 | n/a                                             | -9.07%                      |
|         | N2  | 168.9471                         | 11.53365                  | 126.766191                      | 2.103003 | n/a                                             | -24.97%                     |
|         | O2  | 595.2628                         | 34.40819                  | 514.335227                      | 2.711246 | n/a                                             | -13.60%                     |
|         | H2  | 1849.185                         | 70.50671                  | 1645.7581                       | 3.21636  | n/a                                             | -11.00%                     |
| P-RF-D  | CH4 | 184.9357                         | 21.46799                  | 118.081848                      | 2.072183 | n/a                                             | -36.15%                     |
|         | CO2 | 2538.926                         | 218.0743                  | 2123.91166                      | 3.237136 | n/a                                             | -16.35%                     |
|         | N2  | 130.5626                         | 14.98025                  | 92.1239731                      | 1.964372 | n/a                                             | -29.44%                     |
|         | O2  | 606.7425                         | 35.17128                  | 413.90048                       | 2.616895 | n/a                                             | -31.78%                     |
|         | H2  | 1893.513                         | 73.42637                  | 1791.57503                      | 3.253235 | n/a                                             | -5.38%                      |
Fig. S13. Microstructure analysis of the eight selected candidate polymer models, with the benchmark of PIM-1 shown for comparison. (a) Three-dimensional molecular structures of monomers used to form crosslinked polymers. (b) Top row shows the free-volume elements, including interconnected and disconnected voids, with respect to a probe radius of 0.80 Å. Bottom row shows voids colored with respect to the largest radius probe that can be inserted. Here, only those voids having a minimum radius of 1.42 Å (the kinetic radius of H₂) are shown. (c) Pore size distribution of the polymers, calculated by averaging across five models for each polymer. (d) The calculated density values of the polymers. P-DNN-C1, P-DNN-C2, P-DNN-D1, and P-DNN-D2 have relatively higher densities, while P-RF-C and P-RF-D, P-DNN-C3 and P-DNN-C4 have relatively lower densities compared with PIM-1. For void analysis, we apply the Poreblazer code (https://github.com/SarkisovGroup/PoreBlazer) to directly quantify pore size, and its distribution and interconnectivity, which is based on the Hoshen-Kopelman cluster labeling algorithm. The radius of the probe is chosen to be 1.42 Å, representing the mean Van der Waals radius of hydrogen gas molecules.
| Polymer                  | Model | Simulation value (FFV) | Experiment or other simulations value (FFV) | Ref |
|-------------------------|-------|------------------------|--------------------------------------------|-----|
| AOPIM-1 (flexible fluorinated polyimide) | 1     | 0.263388               |                                             |     |
|                         | 2     | 0.266421               |                                             |     |
|                         | 3     | 0.264481               |                                             |     |
|                         | 4     | 0.266                   |                                             |     |
|                         | 5     | 0.266915               | 0.26507                                    | 79  |
|                         | 6     | 0.266927               |                                             |     |
|                         | 7     | 0.263741               |                                             |     |
|                         | 8     | 0.26427                |                                             |     |
|                         | 9     | 0.265235               |                                             |     |
|                         | 10    | 0.263328               |                                             |     |
| Kapton (semi-crystalline polymer) | 1     | 0.4235                 |                                             |     |
|                         | 2     | 0.417864               |                                             |     |
|                         | 3     | 0.413167               |                                             |     |
|                         | 4     | 0.418509               |                                             |     |
|                         | 5     | 0.40949                | 0.418085                                   | 80  |
|                         | 6     | 0.417786               |                                             |     |
|                         | 7     | 0.416903               |                                             |     |
|                         | 8     | 0.423771               |                                             |     |
|                         | 9     | 0.419132               |                                             |     |
|                         | 10    | 0.420725               |                                             |     |

Table S17. Comparison of experimentally determined fraction of free volumes (FFV) with those of the atomistic in-silico models, generated using the method in this work. Two polymers with small pore sizes are considered: AOPIM-1 (flexible fluorinated polyimide) and Kapton (semi-crystalline polymer).
Fig. S14. Training model verification for the ensemble of 16 DNNs trained on MFFs and gas permeabilities imputed using BLR.
In silico crosslinking of ladder polymers and polyimides

The multi-step cross-linking reaction for ladder polymers is exemplified by the case of PIM-1 (Fig. S15(a)). After 45 of each of the two reactants are packed into a 3D-periodic amorphous cell, there are 180 potential reaction sites in the total system (each monomer has four potential reaction sites). As we increase the cutoff distance from 4.5 to 9.5 Å step by step, the cross-linking degree increases from 33.3% to 92.8% (Fig. S15(b)). The final cross-linked system after relaxation contains 2501 atoms with a density of 1.07 g/cm³, and the box length is 31.87 Å (Fig. S15(c)). During the cross-linking process, extra hydrogen atoms are removed, and partial charges are updated to satisfy charge-neutrality. Then the generated cross-linked polymer structure is used for subsequent calculations of solubility and diffusivity. The ladder polymers in Dataset D of our work are subject to a similar cross-linking procedure, before further MD simulations.

The multi-step crosslinking strategy used for PIM-1 and other ladder polymers is also applied to novel polyimides, namely P-DNN-C1 through P-DNN-C4, whose case studies are shown in Fig. S16-S19, respectively. The difference for polyimides, compared to ladder polymers, is that after 45 of each component are packed into the 3D-periodic amorphous cell, there are 90 potential reaction sites in the total system (each monomer has two potential reaction sites). Differently, the reactive atoms for polyimides are carbon and nitrogen.
Fig. S15. Cross-linking steps for the generation of an atomistic model of PIM-1. (a) The reaction mechanism between 5,5',6,6'-tetrahydroxy-3,3',3'-tetramethyl-1,1'-spirobisindane (PubChem CID 66162) and Tetrafluoroterephthalonitrile (PubChem CID 104426). Carbon-oxygen bonds (C-O) are formed through the crosslinking of highlighted carbon and oxygen accompanied by the elimination of hydrogen molecules. (b) The cross-linking degree increases as the cutoff distance increases from 4.5 to 9.5 Å. The final cross-linking degree reaches 92.8%. (c) Snapshot of the final cross-linked network of PIM-1. Cross-linked C and O atoms are highlighted with sticks and balls, while other atoms are drawn as lines.
Fig. S16. Cross-linking steps for the generation of an atomistic model of P-DNN-C1. (a) The reaction mechanism between the diamine component ‘4-[2-[3-[2-(4-aminocyclohexyl)propan-2-yl]phenyl]propan-2-yl]cyclohexan-1-amine’ (PubChem CID 54351456) and the dianhydride component ‘1,6-Bis(Heptadecafluorooctyl)-3,4,9,10-Perylenetetracarboxylic Acid 3,4:9,10-Dianhydride’ (PubChem CID 46217647). The imide groups are formed through the crosslinking of highlighted carbons and nitrogens, accompanied by eliminating hydrogen molecules. (b) The cross-linking degree increases as the cutoff distance is increased from 4.5 to 20 Å. The final cross-linking degree reaches 90.0%. (c) Snapshot of the final cross-linked network of P-DNN-C1. Cross-linked C and N atoms are highlighted with sticks and balls, while other atoms are drawn as lines.
Fig. S17. Cross-linking steps for the generation of an atomistic model of P-DNN-C2. (a) The reaction mechanism between the diamine component ‘4-[4-(4-amino-2,3,3,4-tetramethylpentan-2-yl)-2,2,3,3,5,5,6,6-octamethylcyclohexyl]-2,3,3,4-tetramethylpentan-2-amine’ (PubChem CID 118069048) and the dianhydride component ‘1,6-Bis(Heptadecafluoroctyl)-3,4,9,10-Perylenetetracarboxylic Acid 3,4,9,10-Dianhydride’ (PubChem CID 46217647). The imide groups are formed through the crosslinking of highlighted carbons and nitrogens accompanied by eliminating hydrogen molecules. (b) The cross-linking degree increases as the cutoff distance is increased from 0.5 to 17.5 Å. The final cross-linking degree reaches 90.0%. (c) Snapshot of the final cross-linked network of P-DNN-C2. Cross-linked C and N atoms are highlighted with sticks and balls, while other atoms are drawn as lines.
Fig. S18. Cross-linking steps for the generation of an atomistic model of P-DNN-C3. (a) The reaction mechanism between the diisocyanate component ‘N-[1,17-ditert-butyl-5-ethyl-8-(2-isocyanatoazepane-1-carbonyl)-19,20-dimethyl-4,12-dioxo-18,21,23-trioxa-3,5,8,11,13-pentaza-1,17-disilabicyclo[15.4.3]tetracosan-11-yl]-2-isocyanato-N-methylazepane-1-carboxamide’ (PubChem CID 60139559) and the dianhydride component ‘Dodecahydro-[5,5’-biisobenzofuran]-1,1’,3,3’-tetraone’ (PubChem CID 11077626). The imide groups are formed through the crosslinking of highlighted carbons and nitrogens, accompanied by eliminating hydrogen molecules. (b) The cross-linking degree increases with the increase of the cutoff distance from 4.5 to 15 Å. The final cross-linking degree reaches 90.0%. (c) Snapshot of the final cross-linked network of P-DNN-C3. Cross-linked C and N atoms are highlighted with sticks and balls, while other atoms are drawn as lines.
Fig. S19. Cross-linking steps for the generation of an atomistic model of P-DNN-C4. (a) The reaction mechanism between the diisocyanate component ‘14-Isocyanato-14-(1-isocyanatoethyl)trispiro [10.0.0.1113.1012.1011] tetratetracontane’ (PubChem CID 141382908) and the dianhydride component ‘2,2’-[Ethylenebis(sulfonyl)]bis(acetic acid)dialectic acid dianhydride’ (PubChem CID 101718050). The imide groups are formed through the crosslinking of highlighted carbons and nitrogens, accompanied by eliminating hydrogen molecules. (b) The cross-linking degree increases with the increase of the cutoff distance from 4.5 to 15 Å. The final cross-linking degree reaches 90.0%. (c) Snapshot of the final cross-linked network of P-DNN-C4. Cross-linked C and N atoms are highlighted with sticks and balls, while other atoms are drawn as lines.
REFERENCES AND NOTES

1. S. Basu, A. L. Khan, A. Cano-Odena, C. Liu, I. F. J. Vankelecom, Membrane-based technologies for biogas separations. *Chem. Soc. Rev.* **39**, 750–768 (2010).

2. S. Zhao, P. H. M. Feron, L. Deng, E. Favre, E. Chabanon, S. Yan, J. Hou, V. Chen, H. Qi, Status and progress of membrane contactors in post-combustion carbon capture: A state-of-the-art review of new developments. *J. Membr. Sci.* **511**, 180–206 (2016).

3. Y. Han, W. S. W. Ho, Polymeric membranes for CO$_2$ separation and capture. *J. Membr. Sci.* **628**, 119244 (2021).

4. B. D. Freeman, Basis of permeability/selectivity tradeoff relations in polymeric gas separation membranes. *Macromolecules* **32**, 375–380 (1999).

5. D. F. Sanders, Z. P. Smith, R. Guo, L. M. Robeson, J. E. McGrath, D. R. Paul, B. D. Freeman, Energy-efficient polymeric gas separation membranes for a sustainable future: A review. *Polymer* **54**, 4729–4761 (2013).

6. H. B. Park, J. Kamcev, L. M. Robeson, M. Elimelech, B. D. Freeman, Maximizing the right stuff: The trade-off between membrane permeability and selectivity. *Science* **356**, eaab0530 (2017).

7. L. M. Robeson, The upper bound revisited. *J. Membr. Sci.* **320**, 390–400 (2008).

8. B. Comesaña-Gándara, J. Chen, C. G. Bezzu, M. Carta, I. Rose, M.-C. Ferrari, E. Esposito, A. Fuoco, J. C. Jansen, N. B. McKeown, Redefining the Robeson upper bounds for CO$_2$/CH$_4$ and CO$_2$/N$_2$ separations using a series of ultrapermeable benzotriptycane-based polymers of intrinsic microporosity. *Energ. Environ. Sci.* **12**, 2733–2740 (2019).

9. R. Swaidan, B. Ghanem, I. Pinnau, Fine-tuned intrinsically ultramicroporous polymers redefine the permeability/selectivity upper bounds of membrane-based air and hydrogen separations. *ACS Macro Lett.* **4**, 947–951 (2015).

10. J. R. Werber, C. O. Osuji, M. Elimelech, Materials for next-generation desalination and water purification membranes. *Nat. Rev. Mater.* **1**, 1–15 (2016).
11. S. Wang, X. Li, H. Wu, Z. Tian, Q. Xin, G. He, D. Peng, S. Chen, Y. Yin, Z. Jiang, M. D. Guiver, Advances in high permeability polymer-based membrane materials for CO$_2$ separations. *Energ. Environ. Sci.* **9**, 1863–1890 (2016).

12. T. J. Corrado, Z. Huang, D. Huang, N. Wamble, T. Luo, R. Guo, Pentiptycene-based ladder polymers with configurational free volume for enhanced gas separation performance and physical aging resistance. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2022204118 (2021).

13. H. Sanaeepur, A. Ebadi Amooghin, S. Bandehali, A. Moghadassi, T. Matsuura, B. Van der Bruggen, Polyimides in membrane gas separation: Monomer’s molecular design and structural engineering. *Prog. Polym. Sci.* **91**, 80–125 (2019).

14. J. Wang, Z. Shi, Y. Zang, H. Jia, M. Teraguchi, T. Kaneko, T. Aoki, Macromolecular design for oxygen/nitrogen permselective membranes—Top-performing polymers in 2020. *Polymers* **13**, 3012 (2021).

15. J. Liu, X. Hou, H. B. Park, H. Lin, High-performance polymers for membrane CO$_2$/N$_2$ separation. *Chem. A Eur. J.* **22**, 15980–15990 (2016).

16. N. B. McKeown, Polymers of intrinsic microporosity (PIMs). *Polymer* **202**, 122736 (2020).

17. T. Corrado, R. Guo, Macromolecular design strategies toward tailoring free volume in glassy polymers for high performance gas separation membranes. *Mol. Syst. Des. Eng.* **5**, 22–48 (2020).

18. M. F. Jimenez-Solomon, Q. Song, K. E. Jelfs, M. Munoz-Ibanez, A. G. Livingston, Polymer nanofilms with enhanced microporosity by interfacial polymerization. *Nat. Mater.* **15**, 760–767 (2016).

19. B. S. Ghanem, R. Swaidan, E. Litwiller, I. Pinnau, Ultra-microporous triptycne-based polyimide membranes for high-performance gas separation. *Adv. Mater.* **26**, 3688–3692 (2014).

20. B. S. Ghanem, R. Swaidan, X. Ma, E. Litwiller, I. Pinnau, Energy-efficient hydrogen separation by AB-type ladder-polymer molecular sieves. *Adv. Mater.* **26**, 6696–6700 (2014).
21. I. Rose, C. G. Bezzu, M. Carta, B. Comesaña-Gándara, E. Lasseuguette, M. C. Ferrari, P. Bernardo, G. Clarizia, A. Fuoco, J. C. Jansen, K. E. Hart, T. P. Liyana-Arachchi, C. M. Colina, N. B. McKeown, Polymer ultrapermeability from the inefficient packing of 2D chains. *Nat. Mater.* **16**, 932–937 (2017).

22. R. C. Dutta, S. K. Bhatia, Atomistic investigation of mixed-gas separation in a fluorinated polyimide membrane. *ACS Appl. Polym. Mater.* **1**, 1359–1371 (2019).

23. W. Fang, L. Zhang, J. Jiang, Polymers of intrinsic microporosity for gas permeation: A molecular simulation study. *Mol. Simul.* **36**, 992–1003 (2010).

24. R. M. Venable, A. Krämer, R. W. Pastor, Molecular dynamics simulations of membrane permeability. *Chem. Rev.* **119**, 5954–5997 (2019).

25. S. Yi, B. Ghanem, Y. Liu, I. Pinnau, W. J. Koros, Ultraselective glassy polymer membranes with unprecedented performance for energy-efficient sour gas separation. *Sci. Adv.* **5**, eaaw5459 (2019).

26. L. M. Robeson, C. D. Smith, M. Langsam, A group contribution approach to predict permeability and permselectivity of aromatic polymers. *J. Membr. Sci.* **132**, 33–54 (1997).

27. E. R. Hensema, M. H. V. Mulder, C. A. Smolders, C. A. Smolders, On the mechanism of gas transport in rigid polymer membranes. *J. Appl. Polym. Sci.* **49**, 2081–2090 (1993).

28. M. H. Cohen, D. Turnbull, Molecular transport in liquids and glasses. *J. Chem. Phys.* **31**, 1164–1169 (1959).

29. G. Chen, Z. Shen, A. Iyer, U. F. Ghumman, S. Tang, J. Bi, W. Chen, Y. Li, Machine-learning-assisted de novo design of organic molecules and polymers: Opportunities and challenges. *Polymers* **12**, 163 (2020).

30. D. J. Audus, J. J. de Pablo, Polymer informatics: Opportunities and challenges. *ACS Macro Lett.* **6**, 1078–1082 (2017).

31. L. Tao, G. Chen, Y. Li, Machine learning discovery of high-temperature polymers. *Patterns* **2**, 100225 (2021).
32. G. Chen, L. Tao, Y. Li, Predicting polymers’ glass transition temperature by a chemical language processing model. *Polymers* **13**, 1898 (2021).

33. L. Tao, V. Varshney, Y. Li, Benchmarking machine learning models for polymer informatics: An example of glass transition temperature. *J. Chem. Inf. Model.* **61**, 5395–5413 (2021).

34. S. Wu, Y. Kondo, M. Kakimoto, B. Yang, H. Yamada, I. Kuwajima, G. Lambard, K. Hongo, Y. Xu, J. Shiomi, C. Schick, J. Morikawa, R. Yoshida, Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm. *Npj Comput. Mater.* **5**, 1–11 (2019).

35. A. Mannodi-Kanakkithodi, G. Pilania, T. D. Huan, T. Lookman, R. Ramprasad, Machine learning strategy for accelerated design of polymer dielectrics. *Sci. Rep.* **6**, 20952 (2016).

36. W. Sun, Y. Zheng, K. Yang, Q. Zhang, A. A. Shah, Z. Wu, Y. Sun, L. Feng, D. Chen, Z. Xiao, S. Lu, Y. Li, K. Sun, Machine learning–assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials. *Sci. Adv.* **5**, eaay4275 (2019).

37. R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S. I. Hong, M. Baldo, R. P. Adams, A. Aspuru-Guzik, Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. *Nat. Mater.* **15**, 1120–1127 (2016).

38. C. L. Ritt, M. Liu, T. A. Pham, R. Epsztein, H. J. Kulik, M. Elimelech, Machine learning reveals key ion selectivity mechanisms in polymeric membranes with subnanometer pores. *Sci. Adv.* **8**, eabl5771 (2022).

39. J. W. Barnett, C. R. Bilchak, Y. Wang, B. C. Benicewicz, L. A. Murdock, T. Bereau, S. K. Kumar, Designing exceptional gas-separation polymer membranes using machine learning. *Sci. Adv.* **6**, eaaz4301 (2020).

40. T. Liu, L. Liu, F. Cui, F. Ding, Q. Zhang, Y. Li, Predicting the performance of polyvinylidene fluoride, polyethersulfone and polysulfone filtration membranes using machine learning. *J. Mater. Chem. A* **8**, 21862–21871 (2020).
41. D. Rogers, M. Hahn, Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **50**, 742–754 (2010).

42. S. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions. arXiv:1705.07874 [cs.AI] (22 May 2017).

43. R. Ma, T. Luo, PIIM: A benchmark database for polymer informatics. *J. Chem. Inf. Model.* **60**, 4684–4690 (2020).

44. A. Ghosh, S. K. Sen, S. Banerjee, B. Voit, Solubility improvements in aromatic polyimides by macromolecular engineering. *RSC Adv.* **2**, 5900–5926 (2012).

45. S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E. E. Bolton, PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Res.* **47**, D1102–D1109 (2019).

46. P. M. Budd, K. J. Msayib, C. E. Tattershall, B. S. Ghanem, K. J. Reynolds, N. B. McKeown, D. Fritsch, Gas separation membranes from polymers of intrinsic microporosity. *J. Membr. Sci.* **251**, 263–269 (2005).

47. N. Du, M. D. Guiver, G. P. Robertson, Ladder polymers with intrinsic microporosity and process for production thereof. U.S. Patent 9,371,429 (2016).

48. L. McInnes, J. Healy, J. Melville, UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv:1802.03426 [stat.ML] (9 February 2018).

49. Q. Yuan, M. Longo, A. W. Thornton, N. B. McKeown, B. Comesaña-Gándara, J. C. Jansen, K. E. Jelfs, Imputation of missing gas permeability data for polymer membranes using machine learning. *J. Membr. Sci.* **627**, 119207 (2021).

50. T. H. Kim, W. J. Koros, G. R. Husk, K. C. O’Brien, Relationship between gas separation properties and chemical structure in a series of aromatic polyimides. *J. Membr. Sci.* **37**, 45–62 (1988).
51. Y. Hu, M. Shiotsuki, F. Sanda, B. D. Freeman, T. Masuda, Synthesis and properties of indan-based polyacetylenes that feature the highest gas permeability among all the existing polymers. *Macromolecules* **41**, 8525–8532 (2008).

52. H. Lin, B. D. Freeman, Gas permeation and diffusion in cross-linked poly(ethylene glycol diacrylate). *Macromolecules* **39**, 3568–3580 (2006).

53. P. Ertl, A. Schuffenhauer, Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J. Chem.* **1**, 8 (2009).

54. J. G. Wijmans, R. W. Baker, The solution-diffusion model: A review. *J. Membr. Sci.* **107**, 1–21 (1995).

55. M. Pardakhti, E. Moharreri, D. Wanik, S. L. Suib, R. Srivastava, Machine learning using combined structural and chemical descriptors for prediction of methane adsorption performance of metal organic frameworks (MOFs). *ACS Comb. Sci.* **19**, 640–645 (2017).

56. L. Hirschfeld, K. Swanson, K. Yang, R. Barzilay, C. W. Coley, Uncertainty quantification using neural networks for molecular property prediction. *J. Chem. Inf. Model.* **60**, 3770–3780 (2020).

57. K. Yang, K. Swanson, W. Jin, C. Coley, P. Eiden, H. Gao, A. Guzman-perez, T. Hopper, B. Kelley, M. Mathea, A. Palmer, V. Settels, T. Jaakkola, K. Jensen, R. Barzilay, Analyzing learned molecular representations for property prediction. *J. Chem. Inf. Model.* **59**, 3370–3388 (2019).

58. C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **1**, 206–215 (2019).

59. S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2**, 56–67 (2020).

60. J. Jiménez-Luna, M. Skalic, N. Weskamp, G. Schneider, Coloring molecules with explainable artificial intelligence for preclinical relevance assessment. *J. Chem. Inf. Model.* **61**, 1083–1094 (2021).
61. N. Du, G. P. Robertson, J. Song, I. Pinnau, S. Thomas, M. D. Guiver, Polymers of intrinsic microporosity containing trifluoromethyl and phenylsulfone groups as materials for membrane gas separation. *Macromolecules* **41**, 9656–9662 (2008).

62. Y. He, F. M. Benedetti, S. Lin, C. Liu, Y. Zhao, H.-Z. Ye, T. Van Voorhis, M. G. De Angelis, T. M. Swager, Z. P. Smith, Polymers with side chain porosity for ultrapermeable and plasticization resistant materials for gas separations. *Adv. Mater.* **31**, 1807871 (2019).

63. Y. Zhao, Y. He, T. M. Swager, Porous organic polymers via ring opening metathesis polymerization. *ACS Macro Lett.* **7**, 300–304 (2018).

64. R. Swaidan, B. Ghanem, E. Litwiller, I. Pinnau, Physical aging, plasticization and their effects on gas permeation in “rigid” polymers of intrinsic microporosity. *Macromolecules* **48**, 6553–6561 (2015).

65. G. Landrum, *RDKit: A Software Suite for Cheminformatics, Computational Chemistry, and Predictive Modeling* (Academic Press, 2013).

66. S. Plimpton, Fast parallel algorithms for short–range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).

67. H. Sun, S. J. Mumby, J. R. Maple, A. T. Hagler, An ab initio CFF93 all-atom force field for polycarbonates. *J. Am. Chem. Soc.* **116**, 2978–2987 (1994).

68. V. Varshney, S. S. Patnaik, A. K. Roy, B. L. Farmer, A molecular dynamics study of epoxy-based networks: Cross-linking procedure and prediction of molecular and material properties. *Macromolecules* **41**, 6837–6842 (2008).

69. L. H. Hall, B. Mohney, L. B. Kier, The electrotopological state: Structure information at the atomic level for molecular graphs. *J. Chem. Inf. Comput. Sci.* **31**, 76–82 (1991).

70. M. C. Sorkun, J. M. V. A. Koelman, S. Er, Pushing the limits of solubility prediction via quality-oriented data selection. *iScience* **24**, 101961 (2021).
71. P. G. Francoeur, D. R. Koes, SolTranNet—A machine learning tool for fast aqueous solubility prediction. *J. Chem. Inf. Model.* **61**, 2530–2536 (2021).

72. A. Chandrasekaran, C. Kim, S. Venkatram, R. Ramprasad, A deep learning solvent-selection paradigm powered by a massive solvent/nonsolvent database for polymers. *Macromolecules* **53**, 4764–4769 (2020).

73. H. Sun, COMPASS: An ab initio force-field optimized for condensed-phase applications overview with details on alkane and benzene compounds. *J. Phys. Chem. B* **102**, 7338–7364 (1998).

74. S.-H. Park, K.-J. Kim, W.-W. So, S.-J. Moon, S.-B. Lee, Gas separation properties of 6FDA-based polyimide membranes with a polar group. *Macromol. Res.* **11**, 157–162 (2003).

75. C. G. Bezzu, M. Carta, M.-C. Ferrari, J. C. Jansen, M. Monteleone, E. Esposito, A. Fuoco, K. Hart, T. P. Liyana-Arachchi, C. M. Colina, N. B. McKeown, The synthesis, chain-packing simulation and long-term gas permeability of highly selective spirobifluorene-based polymers of intrinsic microporosity. *J. Mater. Chem. A* **6**, 10507–10514 (2018).

76. B. Satilmis, M. Lanč, A. Fuoco, C. Rizzuto, E. Tocci, P. Bernardo, G. Clarizia, E. Esposito, M. Monteleone, M. Dendisová, K. Friess, P. M. Budd, J. C. Jansen, Temperature and pressure dependence of gas permeation in amine-modified PIM-1. *J. Membr. Sci.* **555**, 483–496 (2018).

77. B. S. Ghanem, N. B. McKeown, P. M. Budd, J. D. Selbie, D. Fritsch, High-performance membranes from polyimides with intrinsic microporosity. *Adv. Mater. Deerfield Beach Fla.* **20**, 2766–2771 (2008).

78. M. Carta, P. Bernardo, G. Clarizia, J. C. Jansen, N. B. McKeown, Gas permeability of hexaphenylbenzene based polymers of intrinsic microporosity. *Macromolecules* **47**, 8320–8327 (2014).

79. J. Wu, S. Japip, T.-S. Chung, Infiltrating molecular gatekeepers with coexisting molecular solubility and 3D-intrinsic porosity into a microporous polymer scaffold for gas separation. *J. Mater. Chem. A* **8**, 6196–6209 (2020).
80. Y. Liu, A. Tang, J. Tan, C. Chen, D. Wu, H. Zhang, Structure and gas barrier properties of polyimide containing a rigid planar fluorene moiety and an amide group: Insights from molecular simulations. *ACS Omega* **6**, 4273–4281 (2021).