Supporting Information

Germanium Distributions in Zeolites Derived from Neural Network Potentials†

Indranil Saha,[a] Andreas Erlebach,[a] Petr Nachtigall,[a‡] Christopher J. Heard,[a] and Lukáš Grajciar*[a]

[a] *Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University in Prague, 12483, Czech Republic. E-mail: lukas.grajciar@natur.cuni.cz*
‡ Deceased.

1 Details of the Computational Models

Table S1 Energy [meV/atom] and force [meV/Å] errors of the trained NNP ensemble calculated for the test set of the training database (in-domain errors).

NNP	Energy		Forces	
	MAE	RMSE	MAE	RMSE
1	1.51	4.38	31.22	46.63
2	1.57	4.19	32.41	48.38
3	1.53	4.68	32.02	47.53
4	1.44	3.49	31.33	46.55
5	1.62	4.18	36.70	58.18
6	1.65	3.96	35.86	56.61

Table S2 NNP energy [meV/atom] and force [meV/Å] errors calculated for the test set of structures sub-sampled from BHMC runs (out-of-domain errors).

Test cases	Energy		Forces	
	MAE	RMSE	MAE	RMSE
UTL	0.40	0.49	47.93	61.57
BEC	1.82	2.15	61.87	80.22
UOV	1.19	1.24	52.39	68.47
IWW	2.44	2.49	52.03	68.47
CTH	1.11	1.18	53.41	69.38
Average	1.39	1.67	53.53	69.88

Table S3 The characteristics of the unit cells adopted in simulations.

Zeolites	a,b,c **(Å)**	α,β,γ **(°)**	**Volume (Å^3)**	**No. of D4R**	**No of T$_{sites}$ (Total Atoms)**
UTL	(16.27, 13.98, 12.44)	(90, 104, 64)	2455.82	1	38 (114)
BEC	(12.67, 12.67, 13.36)	(90, 90, 90)	2145.84	2	32(96)
CTH	(14.64, 14.77, 27.28)	(90, 90, 140)	3788.13	2	64(192)
UOV	(12.64, 22.44, 22.44)	(59, 90, 90)	5444.72	3	88(264)
IWW	(12.95, 41.89, 12.61)	(90, 90, 90)	6836.50	4	112(336)

Table S4 The characteristics of the super-cells adopted to probe the finite-size effects.

Zeolites	a,b,c **(Å)**	UC extensions	α,β,γ **(°)**	**Volume (Å^3)**	**No. of D4R**	**No of T$_{sites}$ (Total Atoms)**
UTL	(32.53, 27.97, 12.44)	$a(\times2),b(\times2),c$	(90, 104, 64)	9823.29	4	152(456)
BEC	(25.34, 12.67, 26.73)	$a(\times2),b,c(\times2)$	(90, 90, 90)	8583.34	8	128(384)
CTH	(14.64, 29.54, 27.29)	$a,b(\times2),c$	(90, 90, 140)	7576.25	4	128(384)
IWW	(12.95, 41.89, 25.21)	$a,b,c(\times2)$	(90, 90, 90)	13672.99	8	224(672)

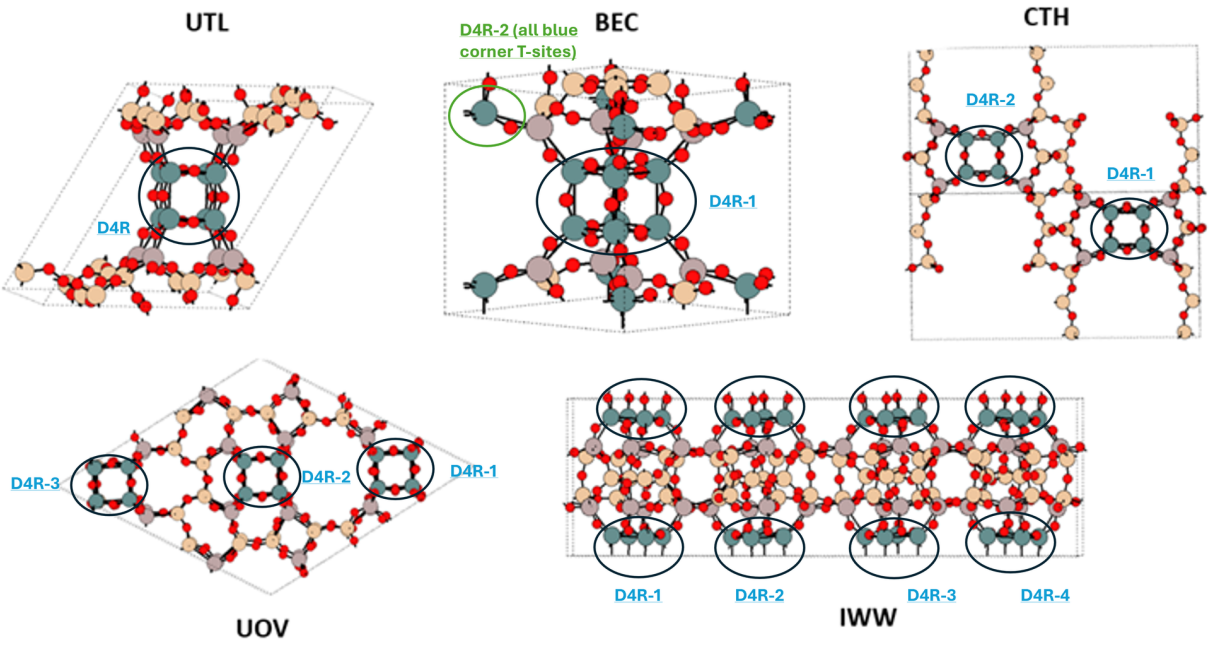


Fig. S1 The structure of the zeolites considered in this work highlighting for each topology three types of T-sites present in the frameworks - the T-sites in blue are the *D4R* sites, the T-sites in violet, directly connected to the D4R, are *adjacent* sites, and the T-sites in yellow are furthest from D4R, called *framework* sites. The D4R units in the unit cell are numbered.

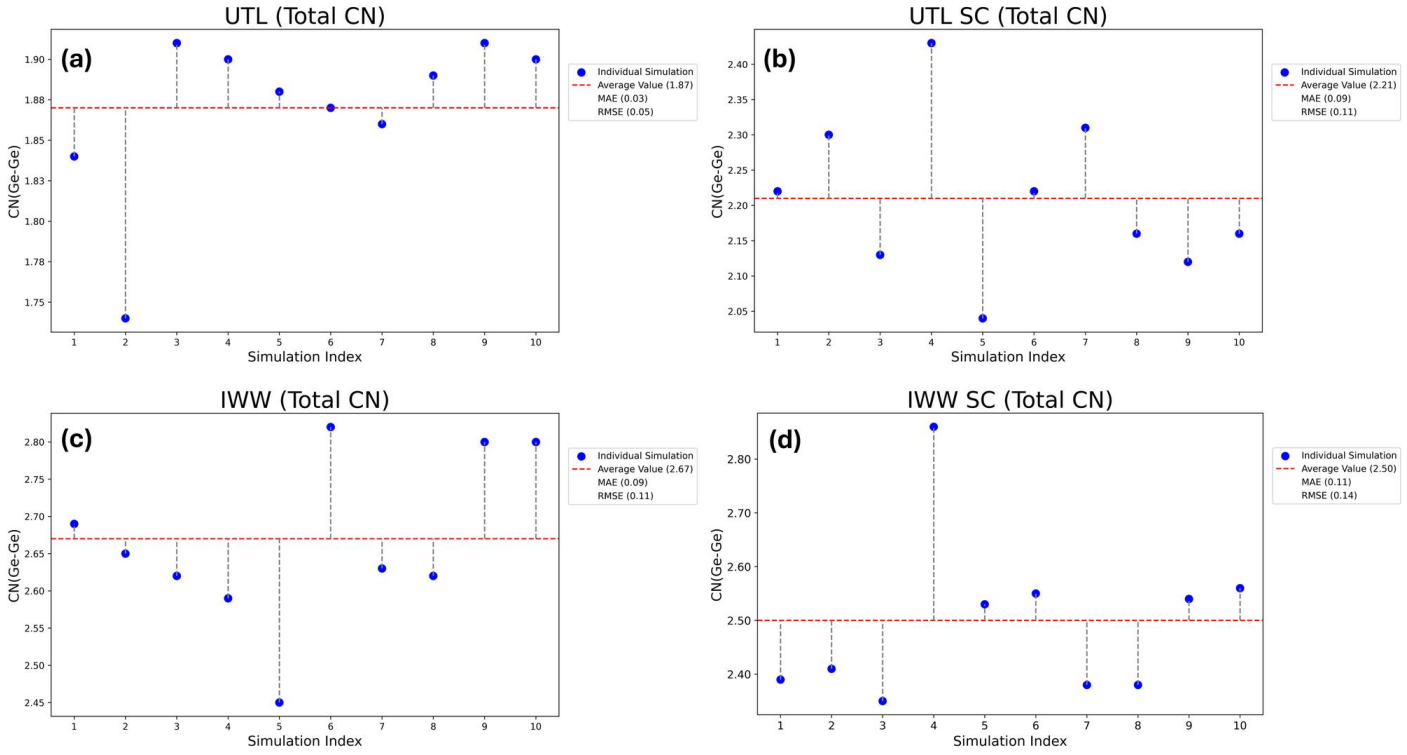


Fig. S2 The evaluation of the statistical uncertainty in the values of the *Ge-Ge coordination number* from 10 independent simulations (in blue, approx. 10K structures considered in each) - the average value (red dashed line) alongside the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) is provided. The data are for the following models,
UTL zeolite with Si:Ge = 2.8 for both (a) single (N_{Ge}=10) and (b) supercell (N_{Ge} = 40) calculations.
IWW zeolite with Si:Ge = 1.5 for both (c) single (N_{Ge}=44) and (d) supercell (N_{Ge} = 88) calculations.

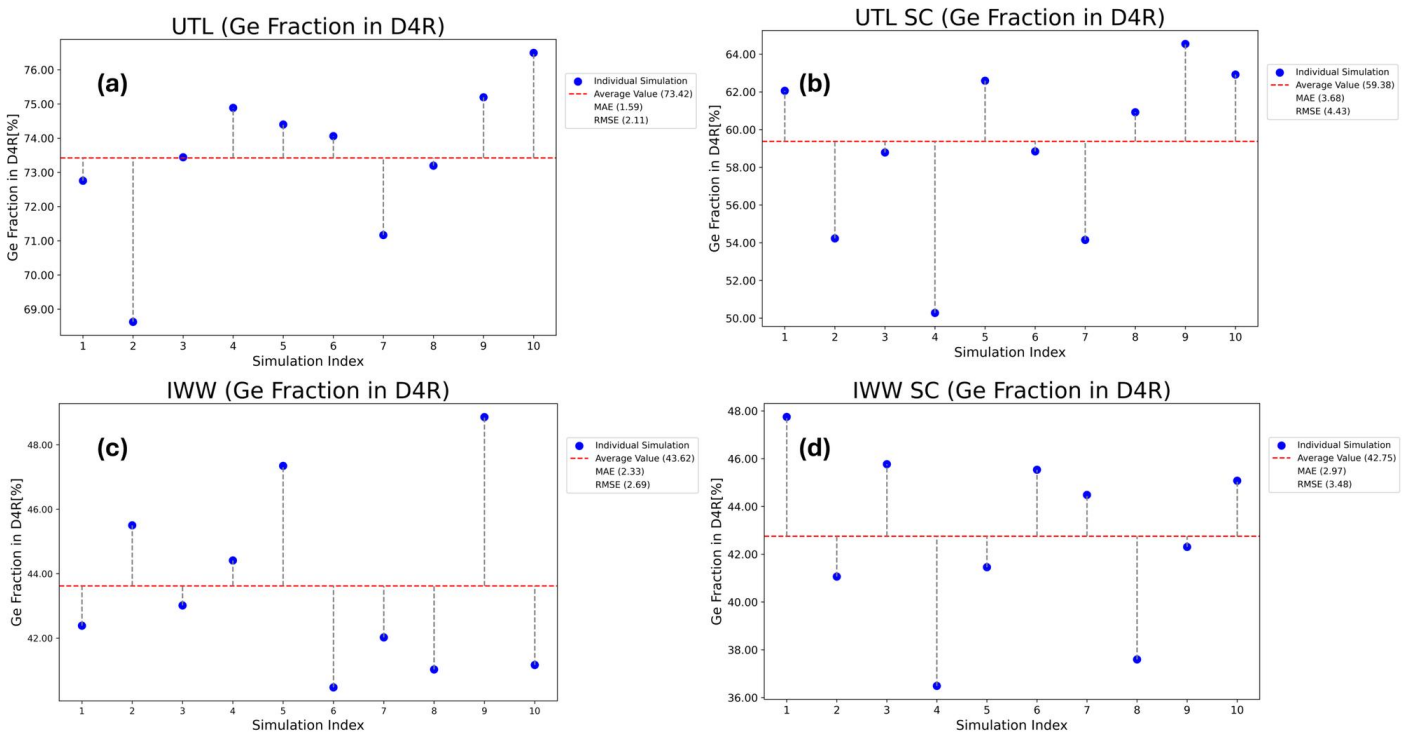


Fig. S3 The evaluation of the statistical uncertainty in the values of the *Ge Fraction in D4R T-sites* from 10 independent simulations (in blue, approx. 10K structures considered in each) - the average value (red dashed line) alongside the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) is provided. The data are for the following models,
UTL zeolite with Si:Ge = 2.8 for both (a)single (N_{Ge}=10) and (b) supercell (N_{Ge} = 40) calculations.
IWW zeolite with Si:Ge = 1.5 for both (c)single (N_{Ge}=44) and (d) supercell (N_{Ge} = 88) calculations.

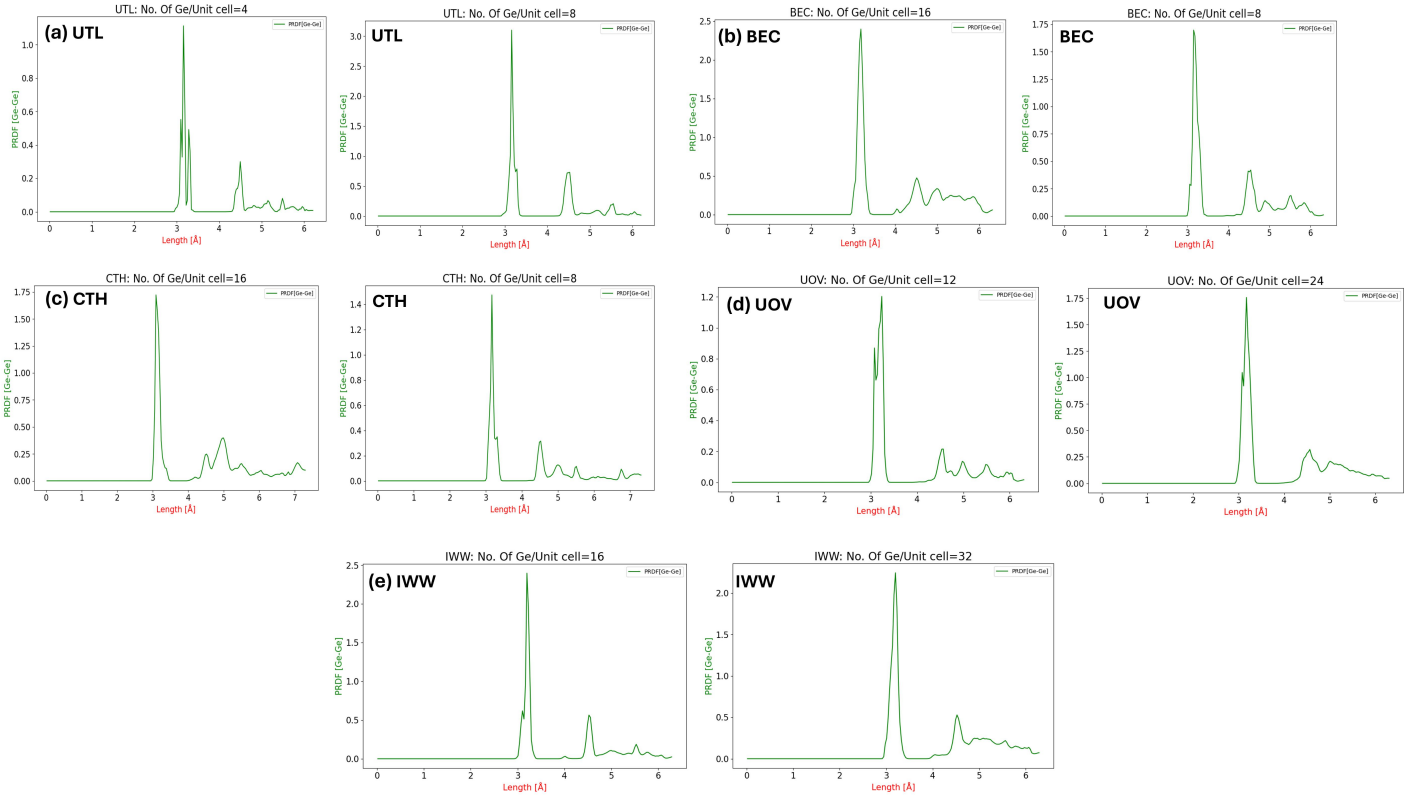


Fig. S4 Partial Radial Distribution Function (PRDF) of Germanium-Germanium (Ge-Ge) distances for all zeolites (a: UTL, b: BEC, c: CTH, d: UOV, e: IWW) and a selected sub-set of Si/Ge ratios corresponding the germanium loading capable of half- and full-filling of the D4R units in the unit cell (considering around 100k structures for each model). Note, that we used a cut-off of 3.5 Å for evaluation of the Ge-Ge coordination number, at which the value(s) of the Ge-Ge PRDFs are zero.

2 Computational Methodology and Structural Optimization of Zeolites

2.1 DFT Setup and DFT-NNP based optimization

To determine suitable computational parameters for performing benchmarking unit cell optimization at the DFT level with the Vienna ab initio simulation package (VASP) using the projector-augmented wave (PAW) pseudopotential, we considered the model of UTL zeolite with Si:Ge= 4.43, in particular its global structure optimum (GSO) obtained via the BHMC simulation (see Section 2.2 for the details on the BHMC set-up. We briefly investigated the influence of various computational parameters (exchange-correlation functional, plane-wave cutoffs, k-point sampling density, dispersion correction type) on the structural properties (see S5). Two functionals of the generalized gradient approximation (GGA) type - Perdew–Burke–Ernzerhof (PBE)[1] revised PBE (revPBE)[2] - were considered. The empirical D3 dispersion correction[3] was considered with two different damping functions - a standard "zero-damping"[4] model and Becke-Johnson (BJ)[5] model. For the standard DFT set-up used throughout this work (PBE-D3(BJ) we examined the effect of three plane-wave energy cutoffs (700, 800, and 900 eV). For all (unit cell) optimization tests, we set the wave function convergence to 10^{-6}eV, the force convergence criterion to 0.01 eV Å^{-1}, and sampled the Brillouin zone with Gamma centred k-point grid.

As shown in Table S5, these tests revealed that all these rather reasonable computational set-up choices have a small (within 3-4 %) but non-negligible effect on the calculated properties of the UTL zeolite. Ultimately, the PBE-D3(BJ) functional at 800 eV was selected as it provided a balance between accuracy and computational efficiency, with the PBE-D3(BJ) functional being used as the reference-level method for generation of the dataset for NNP training (see Section 2.1). This setup was subsequently applied to the DFT-level structure optimization of the GSOs for the remaining zeolites.

Regarding the Neural Network Potential (NNP) based structure optimization, we used our trained potential with the ASE[6] and SchNetPack[7] libraries (see Section 2.1). The BFGS algorithm[8,9] drives the optimization process, ensuring convergence by minimizing the maximum atomic forces to 10^{-3} eV Å^{-1}.

Table S5 Comparison of the effect of exchange-correlation functional, plane-wave cutoff, dispersion correction and k-point grid density (Gamma centred) on unit cell properties of UTL zeolite model with Si:Ge=4.43.

Functional	Cutoff	K-point grid	a,b,c (Å)	α,β,γ (°)	Volume (Å^3)
PBE	**800eV**	**1×1×1**	16.72, 14.17, 12.58	90.06, 103.59, 64.57	2600
PBE-D3(zero)	**800 eV**	**1×1×1**	16.63, 14.08, 12.51	90.19, 103.39, 64.44	2553
revPBE-D3(BJ)	**800 eV**	**1×1×1**	16.55, 13.97, 12.41	90.38, 103.12, 64.31	2503
PBE-D3(BJ)	**700 eV**	**1×1×1**	16.57, 14.01, 12.45	90.28, 103.29, 64.36	2521
PBE-D3(BJ)	**700 eV**	**2×2×2**	16.55, 14.01, 12.46	90.25, 103.21, 64.3521	2521
PBE-D3(BJ)	**700 eV**	**3×3×3**	16.54, 14.0, 12.46	90.27, 103.15, 64.34	2518
PBE-D3(BJ)	**800 eV**	**1×1×1**	16.60, 14.05, 12.48	90.24, 103.36, 64.40	2538
PBE-D3(BJ)	**900 eV**	**1×1×1**	16.61, 14.06, 12.49	90.20, 103.36, 64.40	2545

2.2 Data Acquisition and Analysis

We obtained the GSOs for zeolites with various Si:Ge ratios from the BHMC simulations except for the pure silica and pure germanate forms (for which the germanium distribution is trivial). For each GSO structure, we considered various computational set-ups for unit cell optimisation: i) unit cell optimized using Neural Network Potential (NNP) (denoted as **NNP**$_{GSO}$), ii) unit cell optimized using DFT (denoted as PBE-D3(BJ)) method (**DFT**$_{GSO}$), iii) unit cell optimized using DFT followed by NNP geometry optimisation with the unit cell fixed (denoted as **DFT/NNP**$_{ReoptGSO}$).

These structures were then analysed to procure the data for comparison, i.e., cell parameters, volume, $T-O-T$ (T = Si/Ge) bond angles and $T-O$ (T = Si/Ge) bond lengths (Tables S6 to S15). Also, in the tables for bond angles and bond lengths (Tables S11 to S15), we provide the values for these structural parameters obtained from the BHMC simulation by taking an average across all the BHMC steps, similarly as done in the main text in Section 3.2 (denoted as **NNP**$_{Sim-Avg}$). All the GSO structures, including the optimised GSO using various methods, and the scripts used for the evaluation of the data are available here: https://doi.org/10.5281/zenodo.13357083. Finally, in the Tables S6 to S15 we provide any reported data from the literature (both experimental and from simulation) for the zeolites considered (denoted as **Reported** or **Experiment**).

Table S6 Comparison of cell parameters and unit cell volumes for STW zeolite with various Si/Ge ratios. Data was obtained for different Global Structure Optima (GSO) at their respective Si/Ge ratios: [1] Unit cell optimized using Neural Network Potential (NNP) as NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method as DFT_{GSO}, [3] The experimental data reported in the literature are denoted as Experiment.

STW: Si/Ge		**∞**	**3.0**	**1.5**	**1.1**	**0.7**	**0**
a,b,c **(Å)**	**NNP_{GSO}**	12.05, 12.05, 30.03	12.08, 12.05, 30.01	12.09, 12.09, 30.08	12.13, 12.13, 30.15	12.2, 12.2, 30.41	12.57, 12.57, 30.76
	DFT_{GSO}	12.02, 12.02, 30.06	12.08, 12.05, 30.11	12.11, 12.07, 30.16	12.12, 12.16, 30.15	12.2, 12.2, 30.41	12.69, 12.62, 30.82
	Experiment	11.94, 11.94, 29.75 [10] [11]	12.05, 12.05, 29.95 [10]	12.09, 12.09, 30.08 [10]	12.1, 12.1, 30.08 [10]	12.16, 12.16, 30.25 [10]	12.43, 12.43, 30.63 [10]
Volume(Å^3)	**NNP_{GSO}**	3774	3787	3819	3848.44	3930	4218
	DFT_{GSO}	3763	3798	3829	3848	3930	4278
	Experiment	3650 [10]/3670 [11]	3750 [10]	3810 [10]	3825 [10]	3878 [10]	4096 [10]

Table S7 Comparison of cell parameters and unit cell volumes for BEC zeolite with various Si/Ge ratios. Data was obtained for different Global Structure Optima (GSO) at their respective Si/Ge ratios: [1] Unit cell optimized using Neural Network Potential (NNP) as NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method as DFT_{GSO}, [3] The experimental data reported in the literature are denoted as Experiment.

BEC: Si/Ge		**∞**	**7.0**	**4.3**	**3.57**	**1.91**	**1.0**	**0**
a,b,c **(Å)**	**NNP_{GSO}**	12.56, 12.56, 13.29	12.68, 12.62, 13.38	12.74, 12.65, 13.37	12.78, 12.68, 13.38	12.76, 12.61, 13.38	12.84, 12.69, 13.44	12.82, 12.82, 14.06
	DFT_{GSO}	12.63, 12.63, 13.29	12.75, 12.66, 13.36	12.82, 12.73, 13.32	12.84, 12.78, 13.32	12.76, 12.62, 13.38	12.84, 12.69, 13.45	12.92, 12.92, 14.08
	Experiment	-	-	-	-	12.82, 12.82, 13.35 *Exp* [12]	-	-
Volume(Å^3)	**NNP_{GSO}**	2096	2141	2156	2170	2153	2192	2311
	DFT_{GSO}	2122	2157	2173	2185	2153	2192	2349
	Experiment	-	2145 [13]	2144 [13]	2167 [13]	2194 [12]	2215 [13]	-

Table S8 Comparison of cell parameters and unit cell volumes for AST zeolite with various Si/Ge ratios. Data was obtained for different Global Structure Optima (GSO) at their respective Si/Ge ratios: [1] Unit cell optimized using Neural Network Potential (NNP) as NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method as DFT_{GSO}, [3] The experimental data reported in the literature are denoted as Experiment..

AST: Si/Ge		**∞**	**1.2**	**0.67**	**0.33**	**0.17**	**0**
a,b,c **(Å)**	**NNP_{GSO}**	9.02, 9.02, 13.54	8.72, 8.72, 14.37	8.87, 8.86, 14.33	8.86 8.85, 14.49	8.83, 8.78, 14.67	8.57, 8.54, 15.07
	DFT_{GSO}	9.04, 9.04, 13.82	8.99, 8.99, 14.18	9.11, 9.12, 14.15	9.16, 9.12, 14.24	9.07, 9.14, 14.38	8.73, 8.91, 14.83
	Experiment	9.07, 9.07, 13.44 [14]	9.34, 9.34, 13.75 [15]	9.36, 9.36, 13.89 [15]	9.37, 9.37, 13.99 [15]	9.37, 9.37, 14.04 [15]	9.27, 9.27, 14.35 [14]
Volume(Å^3)	**NNP_{GSO}**	1102	1092	1127	1137	1137	1103
	DFT_{GSO}	1130	1147	1175	1189	1192.21	1154
	Experiment	1105 [14]	1200 [15]	1217 [15]	1229 [15]	1233 [15]	1233 [14]

Table S9 Comparison of cell parameters and unit cell volumes for ASV zeolite with various Si/Ge ratios. Data was obtained for different Global Structure Optima (GSO) at their respective Si/Ge ratios: [1] Unit cell optimized using Neural Network Potential (NNP) as NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method as DFT_{GSO}, [3] The experimental data reported in the literature are denoted as Experiment.

ASV: Si/Ge		**∞**	**4.0**	**1.5**	**0.25**	**0**
a,b,c **(Å)**	**NNP_{GSO}**	8.48, 8.48, 14.18	8.47, 8.47, 14.36	8.57, 8.57, 14.51	8.69, 8.69, 14.89	8.97, 8.97, 15.02
	DFT_{GSO}	8.48, 8.48, 14.12	8.56, 8.56, 14.34	8.66, 8.69, 14.40	8.87, 8.87, 14.65	8.98, 9.07, 14.72
	Experiment	-	-	-	-	8.78, 8.78, 14.47 [16]
Volume(Å^3)	**NNP_{GSO}**	1020	1031	1066	1125	1208
	DFT_{GSO}	1014	1052	1084	1153	1200
	Experiment	-	-	-	-	1115 [17] [16]

Table S10 Comparison of cell parameters and unit cell volumes for UTL zeolite with Si/Ge ratio=4.43. Data was obtained for different Global Structure Optima (GSO) at the respective Si/Ge ratio: [1] Unit cell optimized using Neural Network Potential (NNP) as NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method as DFT_{GSO}, [3] The experimental data reported in the literature are denoted as Experiment.

UTL: Si/Ge		**4.43**
a,b,c **(Å)**	**NNP_{GSO}**	16.31, 14.06, 12.51
	DFT_{GSO}	16.60, 14.05, 12.48
	Experiment	16.46, 13.99, 12.39 [18*]
Volume(Å^3)	**NNP_{GSO}**	2524
	DFT_{GSO}	2538
	Experiment	2494 [18*]

* The unit cell data is adjusted for primitive cell.

Table S11 Comparison of various $T-O-T$ (T = Si/Ge) bond angles and $T-O$ (T = Si/Ge) bond lengths for STW zeolite with different Si/Ge ratios alongside pure germania and silica. Data was obtained using different methods for the average bond angle and lengths: [1] Unit cell optimized using Neural Network Potential (NNP) - NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method - DFT_{GSO}, [3] Unit cell optimized using DFT as before followed by NNP geometry optimisation - $DFT/NNP_{ReoptGSO}$, [4] Average from the BHMC simulation using Neural Network Potential (NNP) - $NNP_{Sim-Avg}$ [5] The data reported in the literature either from theory or experiment based upon availability.

STW: Si/Ge		∞	3.0	1.5	1.1	0.7	0
Ge-O-Ge(°)	NNP_{GSO}		130	130	130	130	132
	DFT_{GSO}		128	128	128	129	130
	$DFT/NNP_{ReoptGSO}$	-	130	130	130	130	133
	$NNP_{Sim-Avg}$		129	129	129	129	129
	Reported		134 [10a]	136 [10a]	136 [10a]	134 [10a]	138 [10a]
Si-O-Ge(°)	NNP_{GSO}		135	135	135	138	
	DFT_{GSO}		134	134	135	137	
	$DFT/NNP_{ReoptGSO}$	-	135	135	135	138	-
	$NNP_{Sim-Avg}$		135	135	135	135	
	Reported		-	-	-	-	
Si-O-Si(°)	NNP_{GSO}	147	146	147	149	151	
	DFT_{GSO}	147	147	149	151	152	
	$DFT/NNP_{ReoptGSO}$	145	146	148	149	151	-
	$NNP_{Sim-Avg}$	-	145	146	147	149	
	Reported	151 [11b]	-	-	-	-	
Si-O(Å)	NNP_{GSO}	1.6246	1.6271	1.6273	1.6275	1.6278	
	DFT_{GSO}	1.6236	1.6257	1.6263	1.6265	1.6270	
	$DFT/NNP_{ReoptGSO}$	1.6267	1.6272	1.6276	1.6275	1.6278	-
	$NNP_{Sim-Avg}$	-	1.6269	1.6265	1.6261	1.6253	
	Reported	1.599 [11a]	-	-		-	
Ge-O(Å)	NNP_{GSO}		1.7562	1.7576	1.7578	1.7584	1.7607
	DFT_{GSO}		1.7797	1.7811	1.7816	1.7818	1.7852
	$DFT/NNP_{ReoptGSO}$	-	1.7563	1.7577	1.7578	1.7584	1.7616
	$NNP_{Sim-Avg}$		1.7534	1.7557	1.7565	1.7575	-
	Reported		1.7549 [11a]	1.7405 [11a]	1.7401 [11a]	1.7337 [a 11]	1.726 [11a]

[a] Theoretical Data.
[b] Experimental Data.

Table S12 Comparison of various $T-O-T$ (T = Si/Ge) bond angles and $T-O$ (T = Si/Ge) bond lengths for BEC zeolite with different Si/Ge ratios alongside pure germania and silica. Data was obtained using different methods for the average bond angle and lengths: [1] Unit cell optimized using Neural Network Potential (NNP) - NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method - DFT_{GSO}, [3] Unit cell optimized using DFT as before followed by NNP geometry optimisation - $DFT/NNP_{ReoptGSO}$, [4] Average from the BHMC simulation using Neural Network Potential (NNP) - $NNP_{Sim-Avg}$ [5] The data reported in the literature either from theory or experiment based upon availability.

BEC: Si/Ge		∞	7.0	4.3	3.57	1.91	1.0	0
Ge-O-Ge(°)	NNP_{GSO}		136	133	132	130	130	133
	DFT_{GSO}		134	131	130	128	128	130
	$DFT/NNP_{ReoptGSO}$	-	136	133	132	130	130	134
	$NNP_{Sim-Avg}$		133	132	132	131	129	-
	Reported		-	-	-	-	-	123.2–144.8 [19a]
Si-O-Ge(°)	NNP_{GSO}		133	134	135	137	139	
	DFT_{GSO}		131	132	134	135	137	
	$DFT/NNP_{ReoptGSO}$	-	134	135	136	137	139	-
	$NNP_{Sim-Avg}$		139	138	138	137	137	
	Reported		-	-	-	-	-	
Si-O-Si(°)	NNP_{GSO}	146	148	148	149	146	147	
	DFT_{GSO}	148	149	149	149	146	148	
	$DFT/NNP_{ReoptGSO}$	147	149	149	149	146	146	-
	$NNP_{Sim-Avg}$	-	147	147	147	146	146	
	Reported	142-162 [20a]	-	-	-	-	-	
Si-O(Å)	NNP_{GSO}	1.6246	1.6243	1.6249	1.6248	1.6269	1.6270	
	DFT_{GSO}	1.6228	1.6231	1.6239	1.6240	1.6261	1.6266	
	$DFT/NNP_{ReoptGSO}$	1.6249	1.6243	1.6249	1.6248	1.6269	1.6270	-
	$NNP_{Sim-Avg}$	-	1.6260	1.6263	1.6264	1.6273	1.6281	
	Reported	-	-	-	-	-	-	
Ge-O(Å)	NNP_{GSO}		1.7549	1.7554	1.7557	1.7567	1.7591	1.7601
	DFT_{GSO}		1.7789	1.7782	1.7782	1.7803	1.7822	1.7849
	$DFT/NNP_{ReoptGSO}$	-	1.7554	1.7560	1.7563	1.7567	1.7591	1.7609
	$NNP_{Sim-Avg}$		1.7492	1.7509	1.7515	1.7533	1.7533	-
	Reported		-	-	-	-	-	1.711-1.807 [19b]

[a] Theoretical Data.
[b] Experimental Data.

Table S13 Comparison of various $T-O-T$ (T = Si/Ge) bond angles and $T-O$ (T = Si/Ge) bond lengths for AST zeolite with different Si/Ge ratios alongside pure germania and silica. Data was obtained using different methods for the average bond angle and lengths: [1] Unit cell optimized using Neural Network Potential (NNP) - NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method - DFT_{GSO}, [3] Unit cell optimized using DFT as before followed by NNP geometry optimisation - $DFT/NNP_{ReoptGSO}$, [4] Average from the BHMC simulation using Neural Network Potential (NNP) - $NNP_{Sim-Avg}$ [5] The data reported in the literature either from theory or experiment based upon availability.

AST: Si/Ge		∞	1.2	0.67	0.33	0.17	0
Ge-O-Ge(°)	NNP_{GSO}		130	132	133	132	130
	DFT_{GSO}		129	130	130	130	128
	$DFT/NNP_{ReoptGSO}$	-	131	133	133	133	131
	$NNP_{Sim-Avg}$		128	128	127	127	-
	Reported		-	-	-	-	133-135 [15b]
Si-O-Ge(°)	NNP_{GSO}		133	134	132	130	130
	DFT_{GSO}		134	134	133	131	128
	$DFT/NNP_{ReoptGSO}$	-	136	136	135	133	131
	$NNP_{Sim-Avg}$		133	132	131	131	-
	Reported		-	-	-	-	
Si-O-Si(°)	NNP_{GSO}	144	142	140	135	-	
	DFT_{GSO}	145	144	141	138	-	
	$DFT/NNP_{ReoptGSO}$	145	144	142	139	-	-
	$NNP_{Sim-Avg}$	-	139	137	135	135	
	Reported	147-153 [20a]	-	-	-	-	
Si-O(Å)	NNP_{GSO}	1.6252	1.6299	1.6320	1.6338	1.6347	
	DFT_{GSO}	1.6233	1.6290	1.6320	1.6356	1.6379	
	$DFT/NNP_{ReoptGSO}$	1.6252	1.6293	1.6312	1.6323	1.6320	-
	$NNP_{Sim-Avg}$	-	1.6286	1.6299	1.6304	1.6309	
	Reported	-	-	-	-	-	
Ge-O(Å)	NNP_{GSO}		1.7574	1.7577	1.7594	1.7604	1.7620
	DFT_{GSO}		1.7789	1.7791	1.7809	1.7828	1.7863
	$DFT/NNP_{ReoptGSO}$	-	1.7571	1.7578	1.7592	1.7602	1.7617
	$NNP_{Sim-Avg}$		1.7522	1.7542	1.7556	1.7565	-
	Reported		-	-	-	-	1.689-1.738 [15b]

[a] Theoretical Data.
[b] Experimental Data.

Table S14 Comparison of various $T-O-T$ (T = Si/Ge) bond angles and $T-O$ (T = Si/Ge) bond lengths for ASV zeolite with different Si/Ge ratios alongside pure germania and silica. Data was obtained using different methods for the average bond angle and lengths: [1] Unit cell optimized using Neural Network Potential (NNP) - NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method - DFT_{GSO}, [3] Unit cell optimized using DFT as before followed by NNP geometry optimisation - $DFT/NNP_{ReoptGSO}$, [4] Average from the BHMC simulation using Neural Network Potential (NNP) - $NNP_{Sim-Avg}$ [5] The data reported in the literature either from theory or experiment based upon availability.

ASV: Si/Ge		∞	4.0	1.5	0.25	0
Ge-O-Ge(°)	NNP_{GSO}		-	128	130	137
	DFT_{GSO}		-	126	128	129
	$DFT/NNP_{ReoptGSO}$	-	-	128	130	132
	$NNP_{Sim-Avg}$		124	126	126	-
	Reported		-	-	-	126-138 [17b]
Si-O-Ge(°)	NNP_{GSO}		130	133	140	
	DFT_{GSO}		128	132	137	
	$DFT/NNP_{ReoptGSO}$	-	130	134	140	-
	$NNP_{Sim-Avg}$		131	132	131	
	Reported		-	-	-	
Si-O-Si(°)	NNP_{GSO}	142	143	144	139	
	DFT_{GSO}	141	147	146	143	
	$DFT/NNP_{ReoptGSO}$	142	144	144	141	-
	$NNP_{Sim-Avg}$	-	141	140	140	
	Reported	134-150 [20a]	-	-	-	
Si-O(Å)	NNP_{GSO}	1.6283	1.6290	1.6297	1.6315	
	DFT_{GSO}	1.6265	1.6269	1.6288	1.6323	
	$DFT/NNP_{ReoptGSO}$	1.6271	1.6298	1.6306	1.6327	-
	$NNP_{Sim-Avg}$	-	1.6285	1.6288	1.6232	
	Reported		-	-	-	
Ge-O(Å)	NNP_{GSO}		1.7573	1.7561	1.7580	1.7524
	DFT_{GSO}		1.7758	1.7791	1.7830	1.7854
	$DFT/NNP_{ReoptGSO}$	-	1.7578	1.7571	1.7587	1.7611
	$NNP_{Sim-Avg}$		1.7509	1.7496	1.7495	-
	Reported		-	-	-	1.735 [15b]

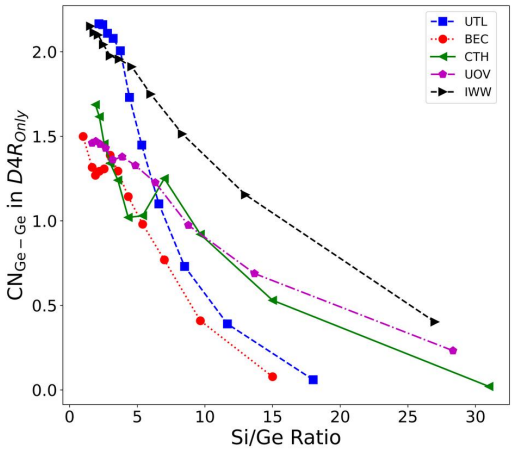
[a] Theoretical Data.
[b] Experimental Data.

Table S15 Comparison of various $T-O-T$ (T = Si/Ge) bond angles and $T-O$ (T = Si/Ge) bond lengths for UTL zeolite with Si/Ge ratio=4.43. Data was obtained using different methods for the average bond angle and lengths: [1] Unit cell optimized using Neural Network Potential (NNP) - NNP_{GSO}, [2] Unit cell optimized using Density Functional Theory [PBE-D3(BJ)] method - DFT_{GSO}, [3] Unit cell optimized using DFT as before followed by NNP geometry optimisation - $DFT/NNP_{ReoptGSO}$, [4] Average from the BHMC simulation using Neural Network Potential (NNP) - $NNP_{Sim-Avg}$ [5] The data reported in the literature either from theory or experiment based upon availability.

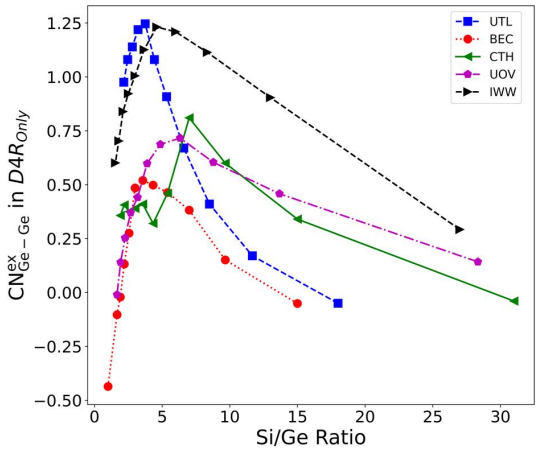
UTL: Si/Ge		**4.43**
Ge-O-Ge(°)	**NNP_{GSO}**	132
	DFT_{GSO}	131
	$DFT/NNP_{ReoptGSO}$	133
	$NNP_{Sim-Avg}$	130
	Reported	128-135[21a]
Si-O-Ge(°)	**NNP_{GSO}**	130
	DFT_{GSO}	129
	$DFT/NNP_{ReoptGSO}$	131
	$NNP_{Sim-Avg}$	133
	Reported	135-155[21a]
Si-O-Si(°)	**NNP_{GSO}**	149
	DFT_{GSO}	150
	$DFT/NNP_{ReoptGSO}$	150
	$NNP_{Sim-Avg}$	147
	Reported	145-162[21a]
Si-O(Å)	**NNP_{GSO}**	1.6238
	DFT_{GSO}	1.6228
	$DFT/NNP_{ReoptGSO}$	1.6329
	$NNP_{Sim-Avg}$	1.6326
	Reported	1.61-1.64[21a]
Ge-O(Å)	**NNP_{GSO}**	1.7583
	DFT_{GSO}	1.7819
	$DFT/NNP_{ReoptGSO}$	1.7598
	$NNP_{Sim-Avg}$	1.7504
	Reported	1.77-1.80[21a]

[a] Theoretical Data.
[b] Experimental Data.
[*] The data doesn't correspond to the particular Si/Ge ratio.[21]

3 Results: Metrics Characterizing Germanium Distributions

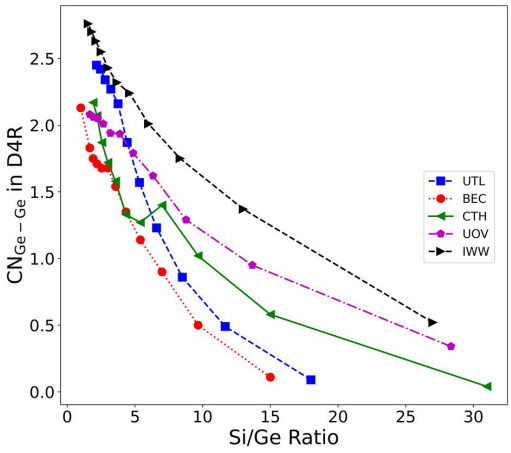


(a) Coordination number in D4R T-sites (for Ge-Ge) [CN_{Ge-Ge}] across zeolites at various Si/Ge ratios

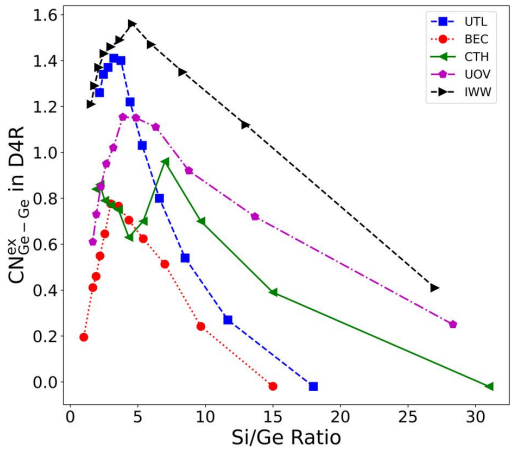


(b) Excess coordination number in D4R T-sites (for Ge-Ge) [CN^{ex}_{Ge-Ge}] across zeolites at various Si/Ge ratios

Fig. S5 Quantification of the clustering tendency of germanium (irrespective of the T site type) across a range of Si/Ge ratios and all zeolite topologies considered focused only on the clustering in D4R T-sites only (See SI Figure S1), namely: (a) CN(Ge-Ge) [CN_{Ge-Ge}]: Ge-Ge coordination number, and (b) Excess CN(Ge-Ge) [CN^{ex}_{Ge-Ge}]: the excess Ge-Ge coordination number, measuring the "over"-clustering of germanium compared to the degree of germanium clustering for a uniform germanium distribution. The CNs are evaluated only for the D4R T-sites considering only the nearest neighbours located in the D4R unit (i.e., not in the adjacent T sites) - this limits the maximum Ge-Ge CN to three only. (See Section 2.3 in the main text)

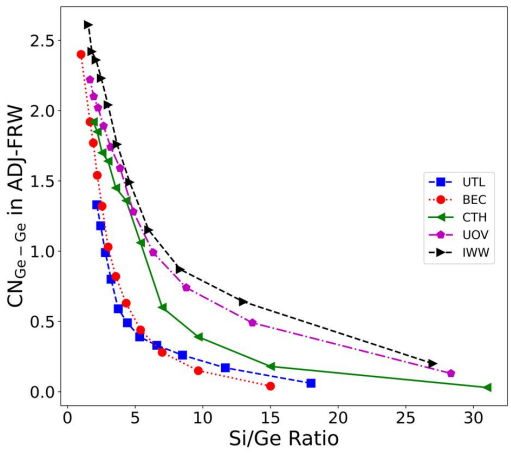


(a) Coordination-Number in D4R T-sites (for Ge-Ge) [CN_{Ge-Ge}] across zeolites at various Si/Ge ratios

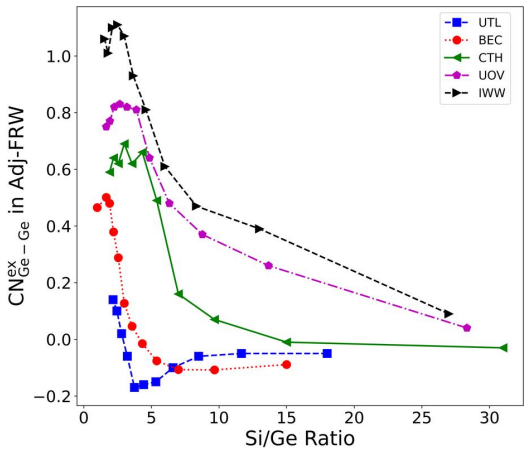


(b) Excess coordination number in D4R T-sites (for Ge-Ge) [CN^{ex}_{Ge-Ge}] across zeolites at various Si/Ge ratios

Fig. S6 Quantification of the clustering tendency of germanium (irrespective of the T site) across a range of Si/Ge ratios and all zeolite topologies considered focused only on the clustering in D4R T-sites, namely: (a) CN(Ge-Ge) [CN_{Ge-Ge}]: Ge-Ge coordination number, and (b) Excess CN(Ge-Ge) [CN^{ex}_{Ge-Ge}]: the excess Ge-Ge coordination number, measuring the "over"-clustering of germanium compared to the degree of germanium clustering for a uniform germanium distribution. The CNs are evaluated only for the D4R T-sites, but now considering all the nearest neighbours including those in the adjacent T sites (See SI Figure S1) - the maximum Ge-Ge CN can be then four. (See Section 2.3 in the main text)



(a) Coordination number (for Ge-Ge) [CN_{Ge-Ge}] in adjacent and framework T-sites across zeolites at various Si/Ge ratios



(b) Excess coordination number in adjacent and framework T-sites (for Ge-Ge) [CN^{ex}_{Ge-Ge}] across zeolites at various Si/Ge ratios

Fig. S7 Quantification of the clustering tendency of germanium (irrespective of the T site) across a range of Si/Ge ratios and all zeolite topologies considered focused only on the clustering in adjacent and framework sites (See SI Figure S1), namely: (a) CN(Ge-Ge) [CN_{Ge-Ge}]: Ge-Ge coordination number, and (b) Excess CN(Ge-Ge) [CN^{ex}_{Ge-Ge}]: the excess Ge-Ge coordination number, measuring the "over"-clustering of germanium compared to the degree of germanium clustering for a uniform germanium distribution. (See Section 2.3 in the main text)

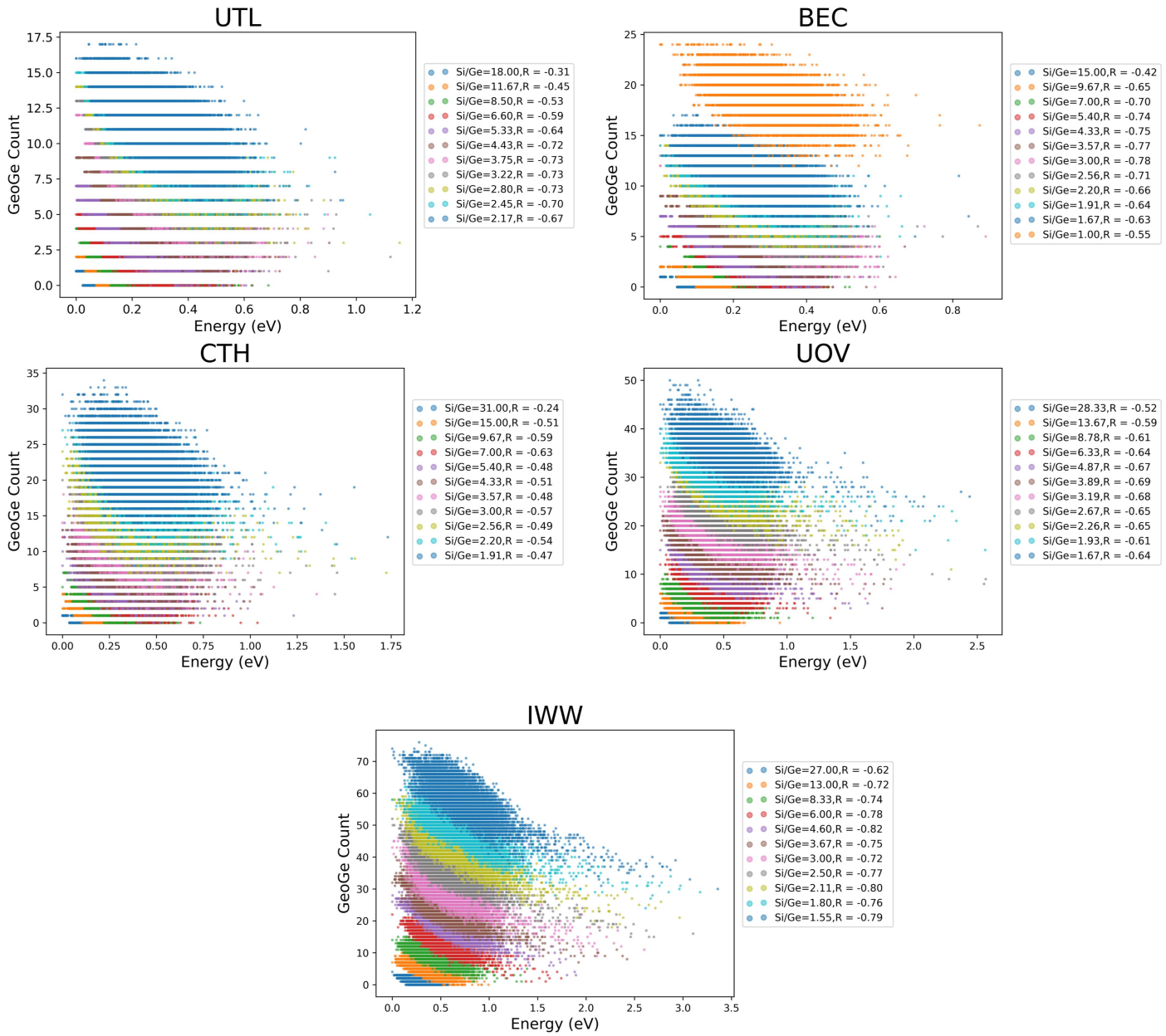


Fig. S8 The correlation between the unique Ge-O-Ge count in a structure and a potential energy of the structure (in eV) for the zeolite topologies (single or unit cells) considered across a range of Si/Ge ratios. The Pearson's correlation coefficients (R)[22] are provided in the legend descriptions.

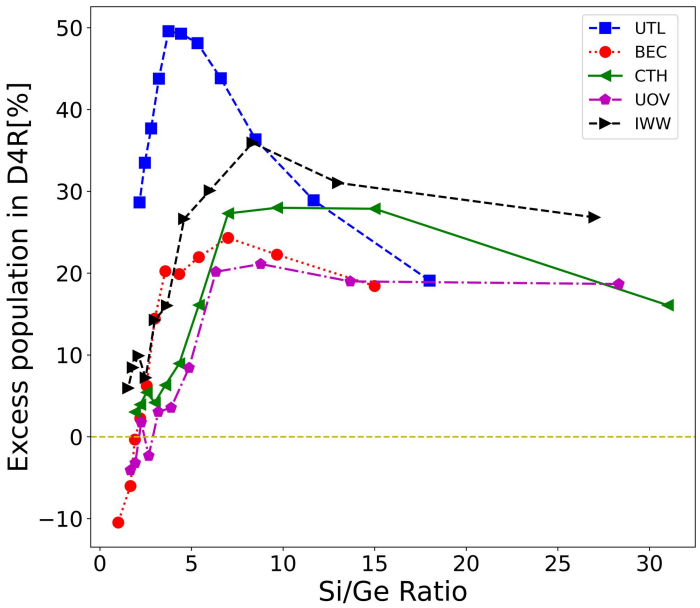


Fig. S9 The Excess Occupation of germanium in D4R [in %], measuring the excess occupation of D4R units by germanium compared to the case of uniform germanium distribution (yellow dashed line) across all the T sites. (See Section 2.3 in the main text)

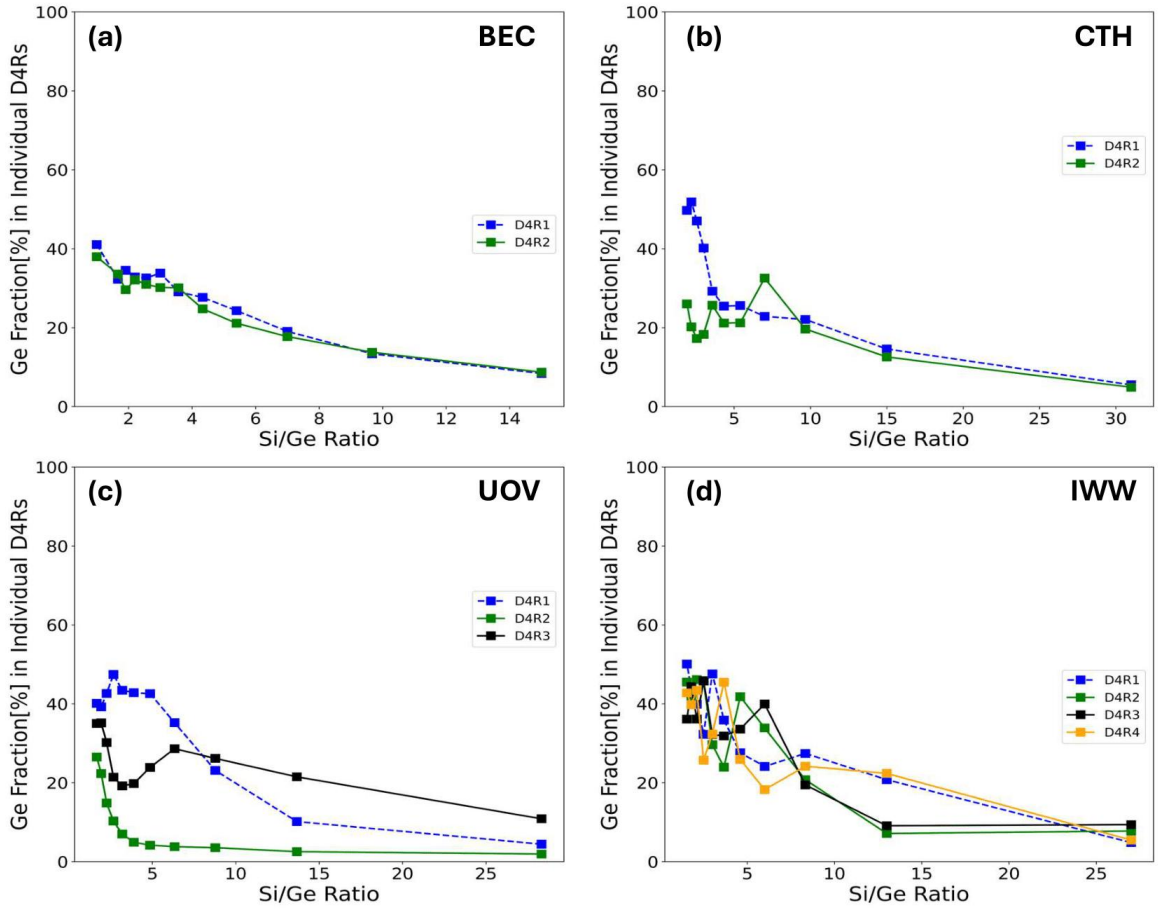


Fig. S10 The Ge Fraction in different D4R sites (See Section 2.3 in the main text) as a function of Si/Ge ratio for the zeolite topologies with more than one D4R unit per unit cell (i.e., except UTL); the numbering of D4R units follows the SI Figure S1, with each of the D4R units being composed of the following crystallographically inequivalent T sites (based on the IZA nomenclature[23]): (a) BEC: D4R1/2-[T1], (b) CTH: D4R1/2-[T2] , (c) UOV: D4R1-[T3-T6] , D4R2-[T3-T6] , D4R3-[T17-19], (d) IWW: D4R1/2/3/4- [T1-T4].

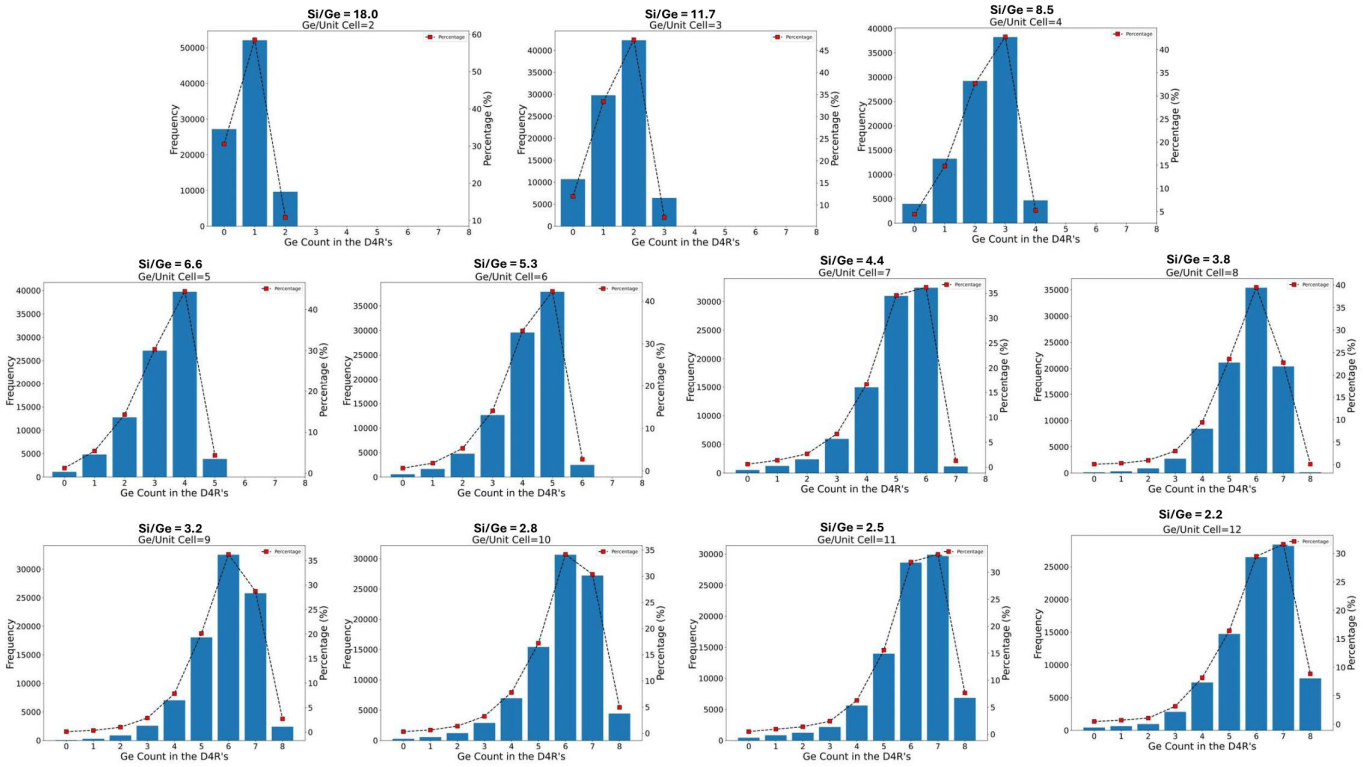


Fig. S11 The frequency distribution of the Ge counts in D4R unit of the UTL zeolite (single-cell model), for approx. 100k structures sampled for each of Si/Ge ratio considered (or alternatively expressed via N_{Ge}/unit-cell). The histogram (bars) shows the frequency of occurrences (primary y-axis) for each Ge count from 0 to 8 (x-axis) in a D4R, while the line plot (purple) represents the percentage of structures (D4Rs) with the corresponding Ge count in the D4R (secondary y-axis) amongst all the structures (D4Rs).

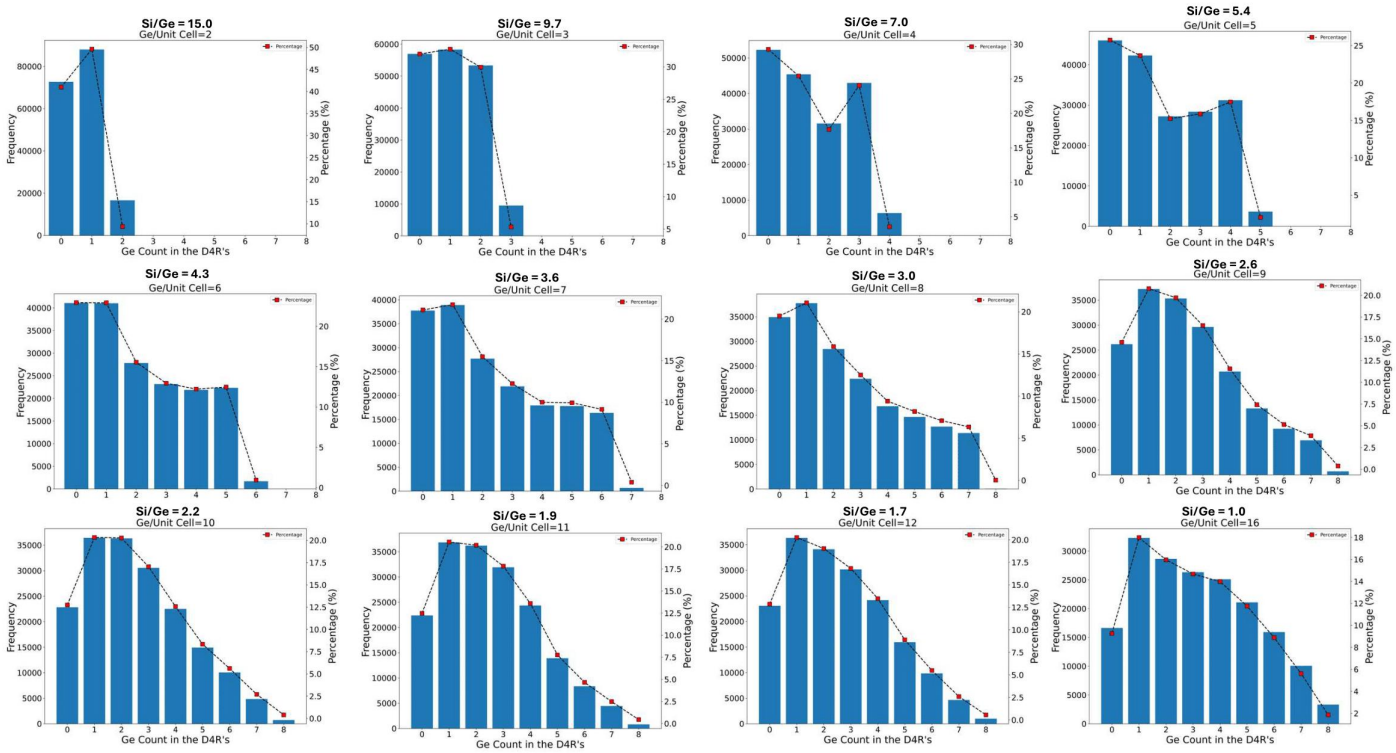


Fig. S12 The frequency distribution of the Ge counts in D4R unit of the BEC zeolite (single-cell model), for approx. 100k structures sampled for each of Si/Ge ratio considered (or alternatively expressed via N_{Ge}/unit-cell). The histogram (bars) shows the frequency of occurrences (primary y-axis) for each Ge count from 0 to 8 (x-axis) in a D4R, while the line plot (purple) represents the percentage of D4Rs with the corresponding Ge count (secondary y-axis) amongst all the structures (D4Rs).

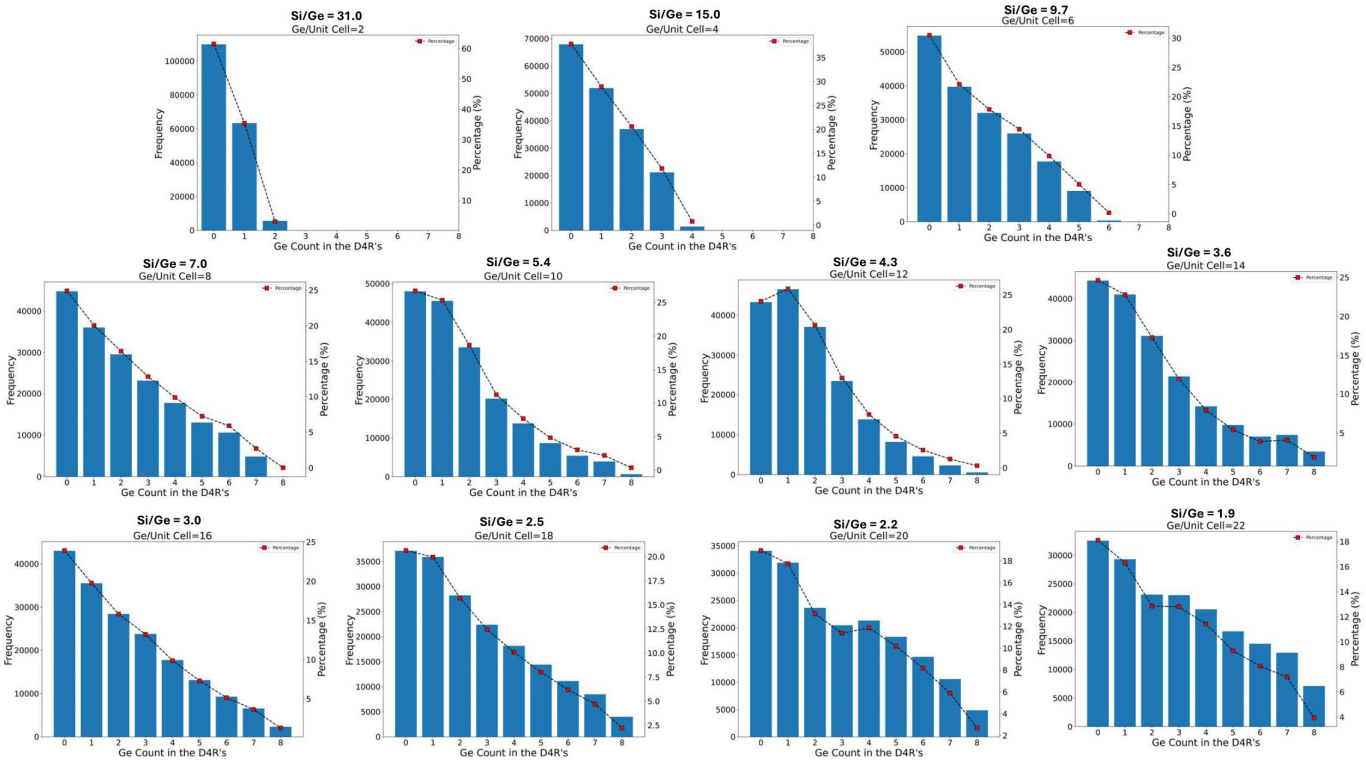


Fig. S13 The frequency distribution of the Ge counts in D4R unit of the CTH(-A) zeolite (single-cell model), for approx. 100k structures sampled for each of Si/Ge ratio considered (or alternatively expressed via N_{Ge}/unit-cell). The histogram (bars) shows the frequency of occurrences (primary y-axis) for each Ge count from 0 to 8 (x-axis) in a D4R, while the line plot (purple) represents the percentage of D4Rs with the corresponding Ge count (secondary y-axis) amongst all the structures (D4Rs).

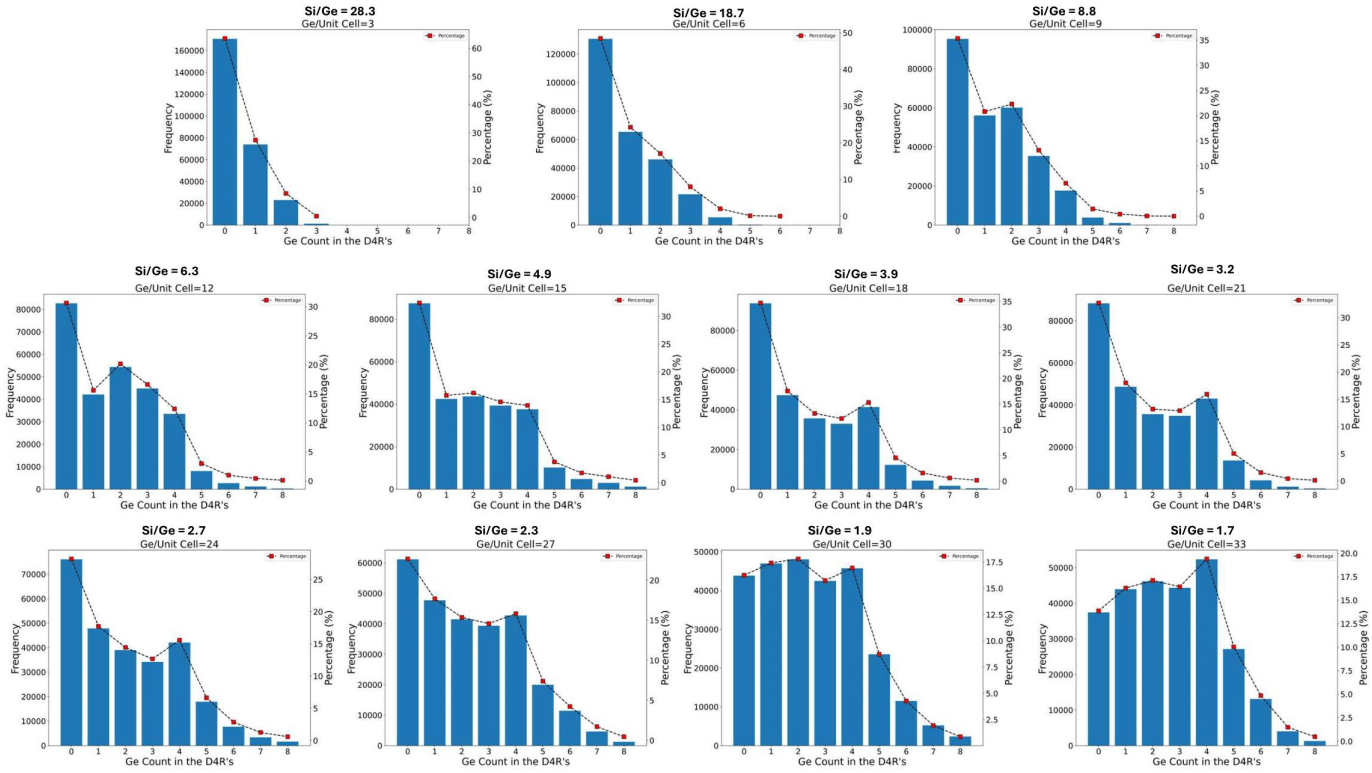


Fig. S14 The frequency distribution of the Ge counts in D4R unit of the UOV zeolite (single-cell model), for approx. 100k structures sampled for each of Si/Ge ratio considered (or alternatively expressed via N_{Ge}/unit-cell). The histogram (bars) shows the frequency of occurrences (primary y-axis) for each Ge count from 0 to 8 (x-axis) in a D4R, while the line plot (purple) represents the percentage of D4Rs with the corresponding Ge count (secondary y-axis) amongst all the structures (D4Rs).

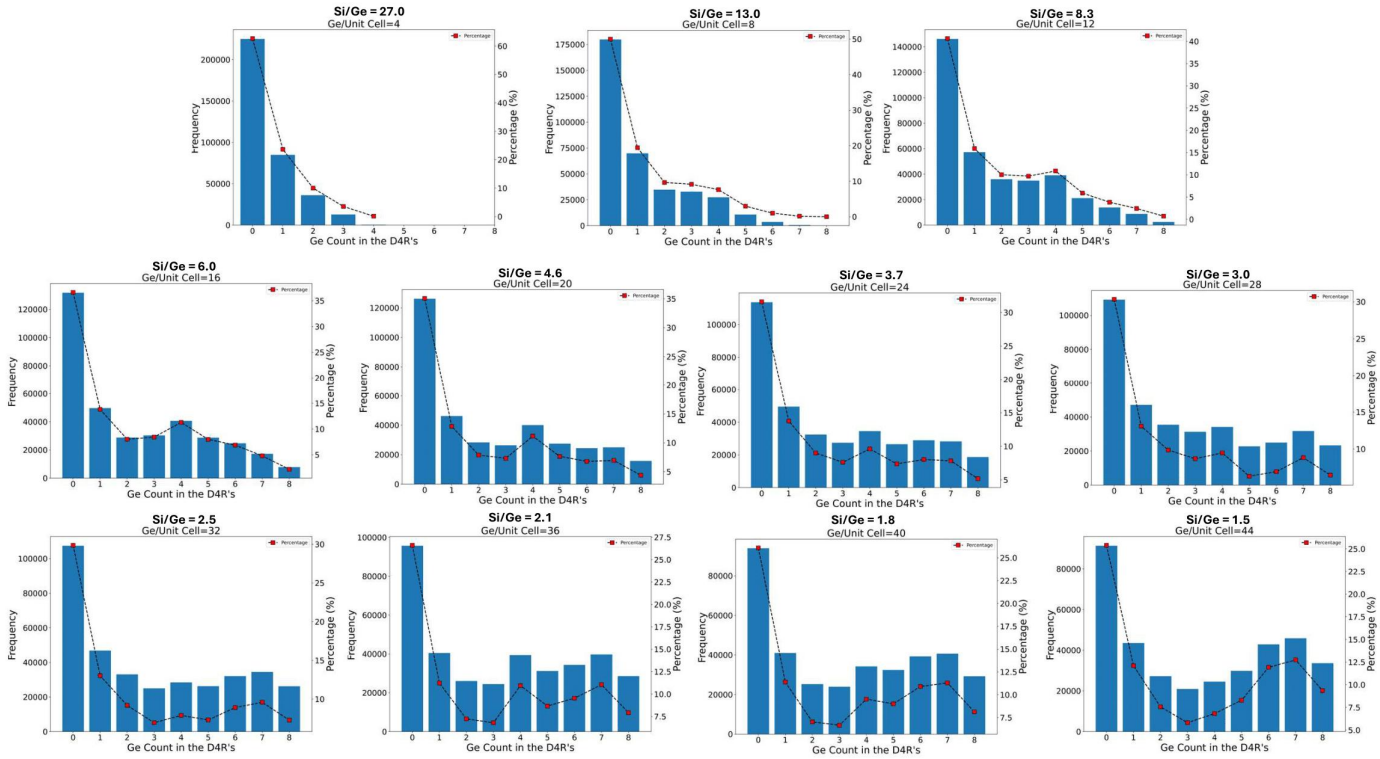


Fig. S15 The frequency distribution of the Ge counts in D4R unit of the IWW zeolite (single-cell model), for approx. 100k structures sampled for each of Si/Ge ratio considered (or alternatively expressed via N_{Ge}/unit-cell). The histogram (bars) shows the frequency of occurrences (primary y-axis) for each Ge count from 0 to 8 (x-axis) in a D4R, while the line plot (purple) represents the percentage of D4Rs with the corresponding Ge count (secondary y-axis) amongst all the structures (D4Rs).

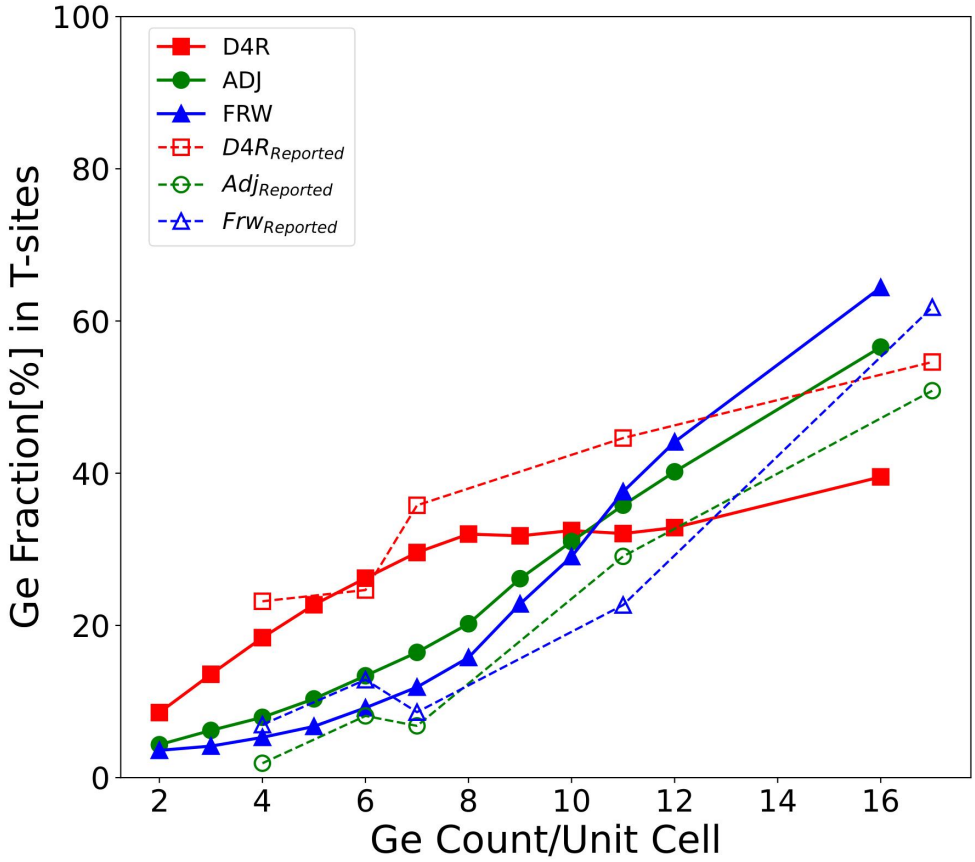


Fig. S16 The Ge fraction in different types of T-sites (D4R, adjacent, framework) in BEC zeolite (single-cell model) at various germanium loadings per unit cell, showing the relative preference of germanium for these T-sites. Solid lines represent our data, while dashed lines indicate reported data from the literature[13] for comparison

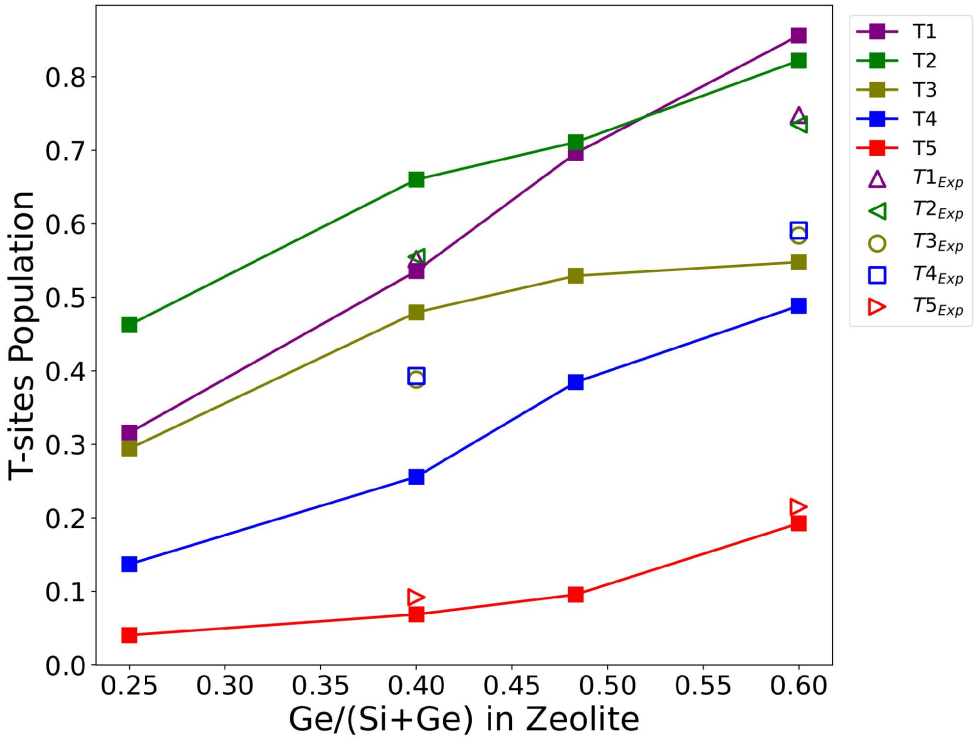


Fig. S17 Germanium occupation of different T-sites (T1-T5) in STW zeolite as a function of germanium content (Ge/(Si + Ge)) in the zeolite. Solid lines with filled markers represent data obtained from our NNP-based BHMC simulations2.2, while scattered empty marker points correspond to experimental data reported in the literature[10] for comparison.

4 Results: Global Structure Optima (GSO) for all zeolites and Si/Ge ratios

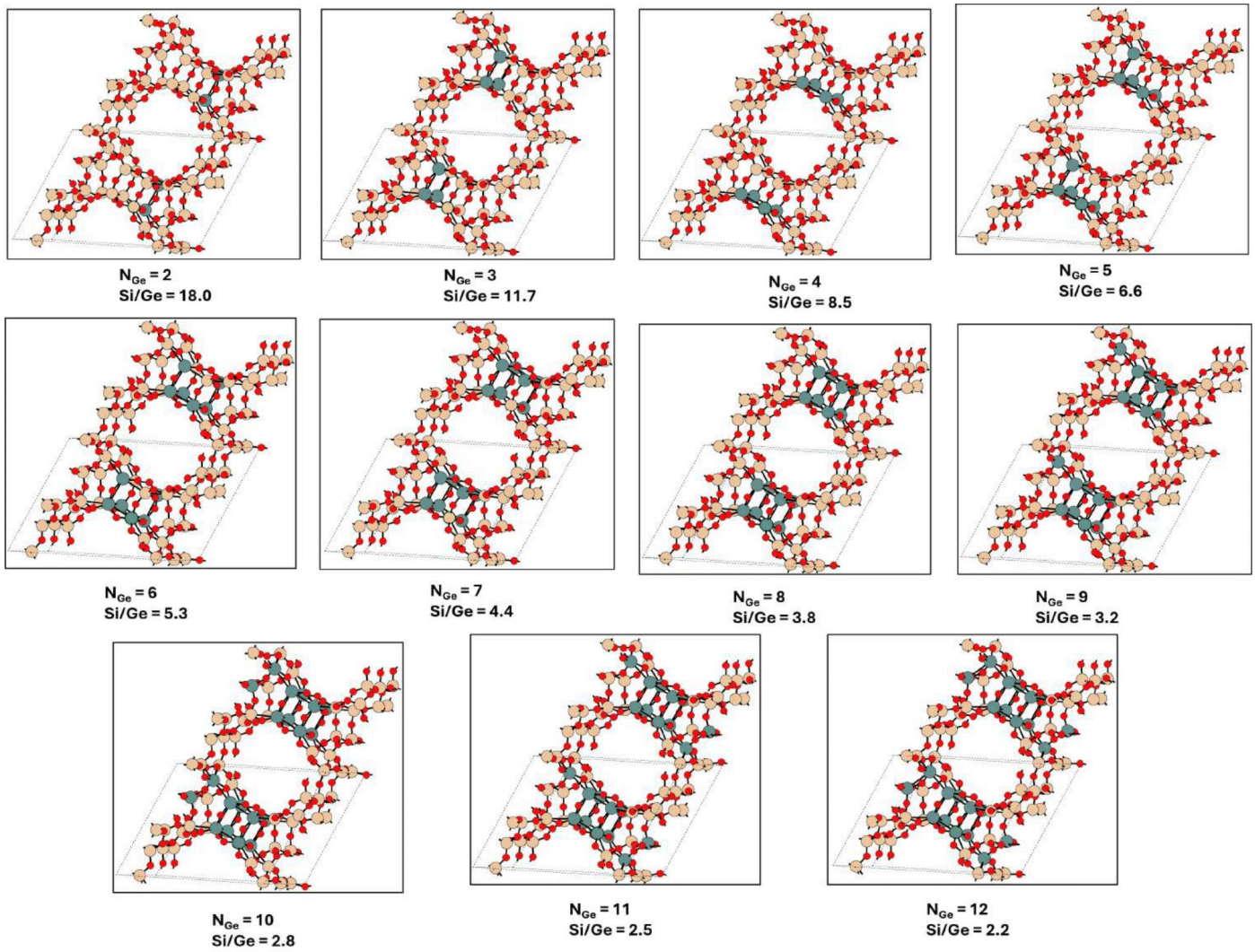


Fig. S18 The global structure optima for UTL zeolite (single-cell model) at all Si/Ge ratios considered in this work, where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

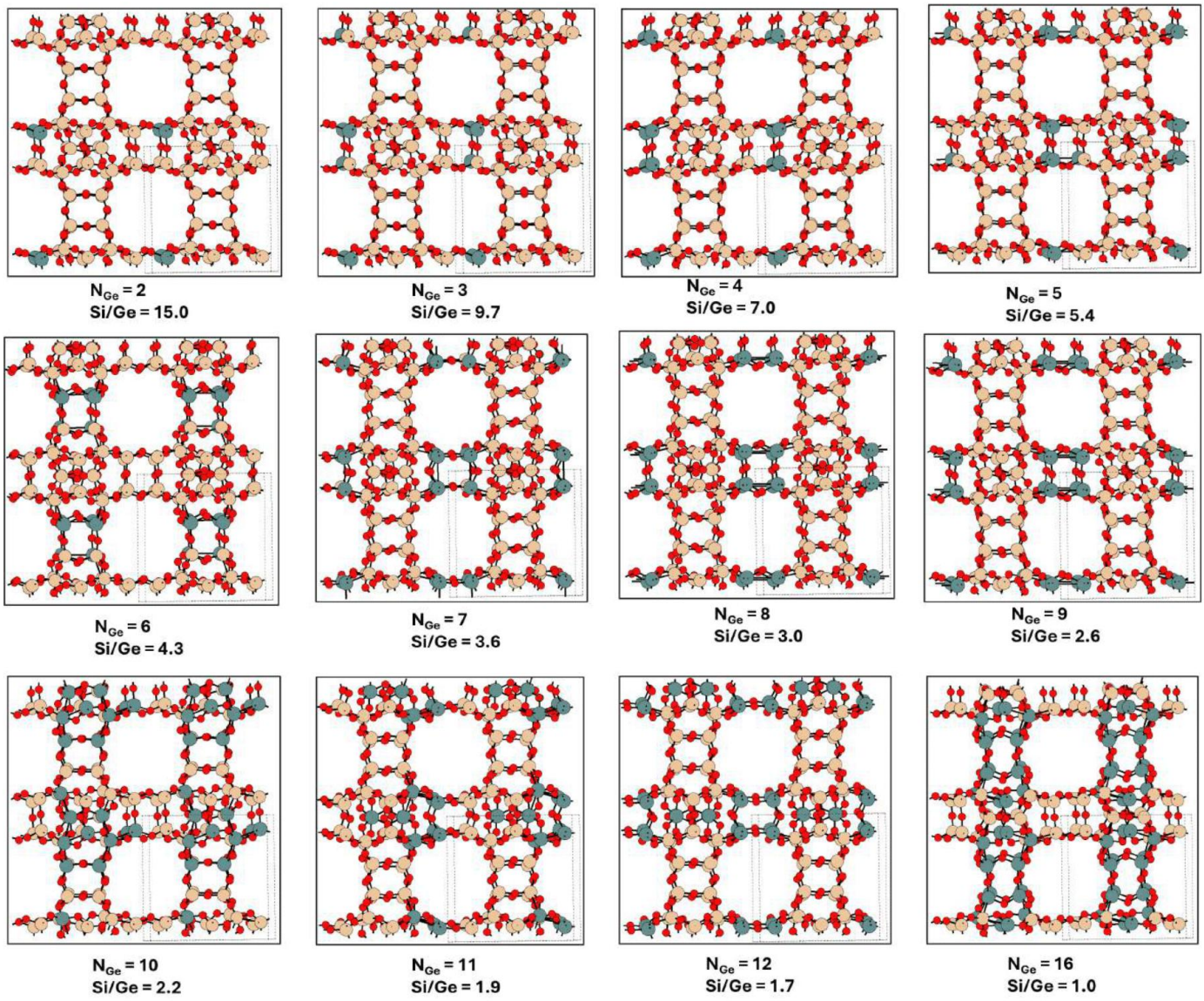


Fig. S19 The global structure optima for BEC zeolite (single-cell model) at all Si/Ge ratios considered in this work, where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

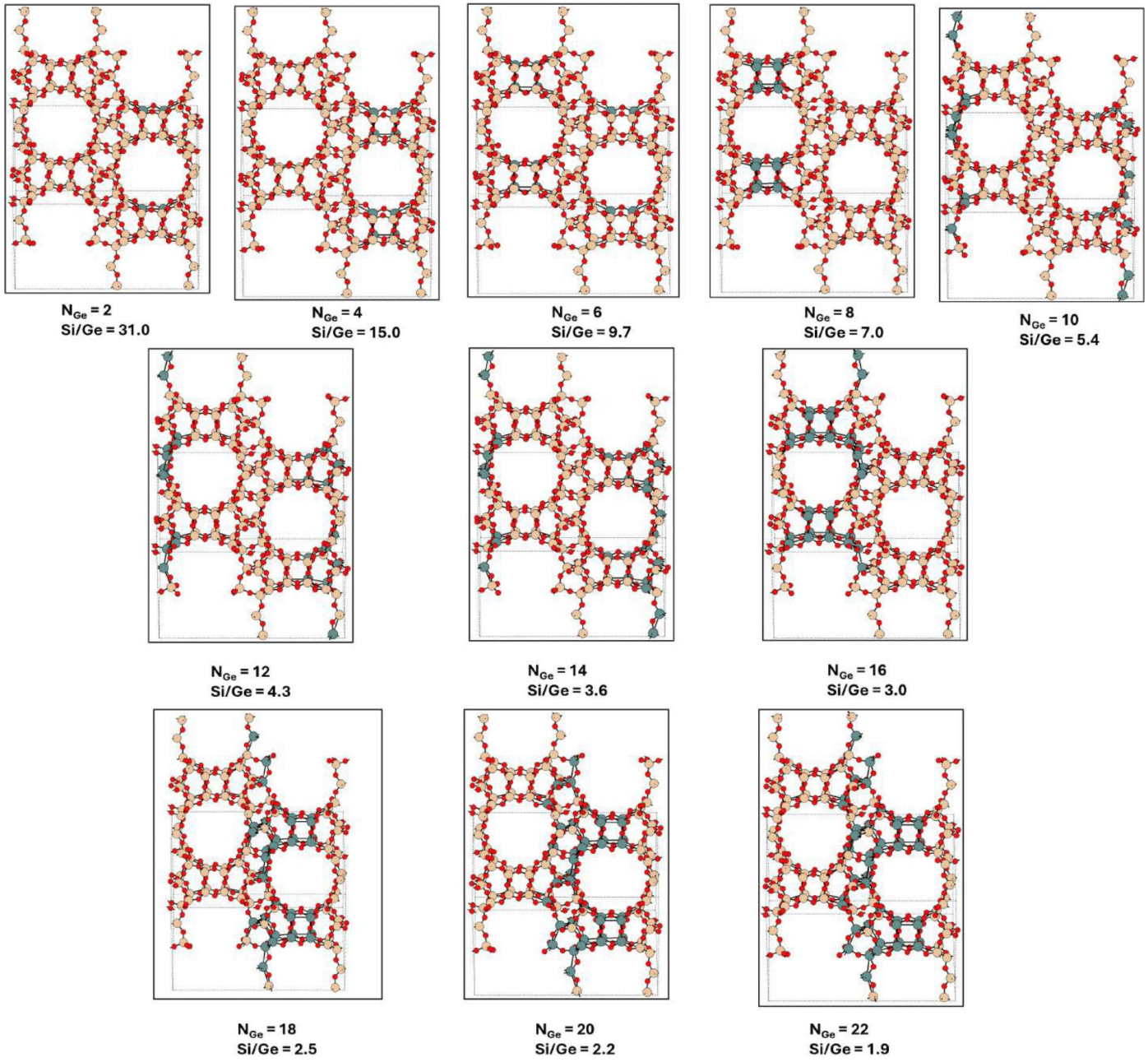


Fig. S20 The global structure optima for CTH(-A) zeolite (single-cell model) at all Si/Ge ratios considered in this work, where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

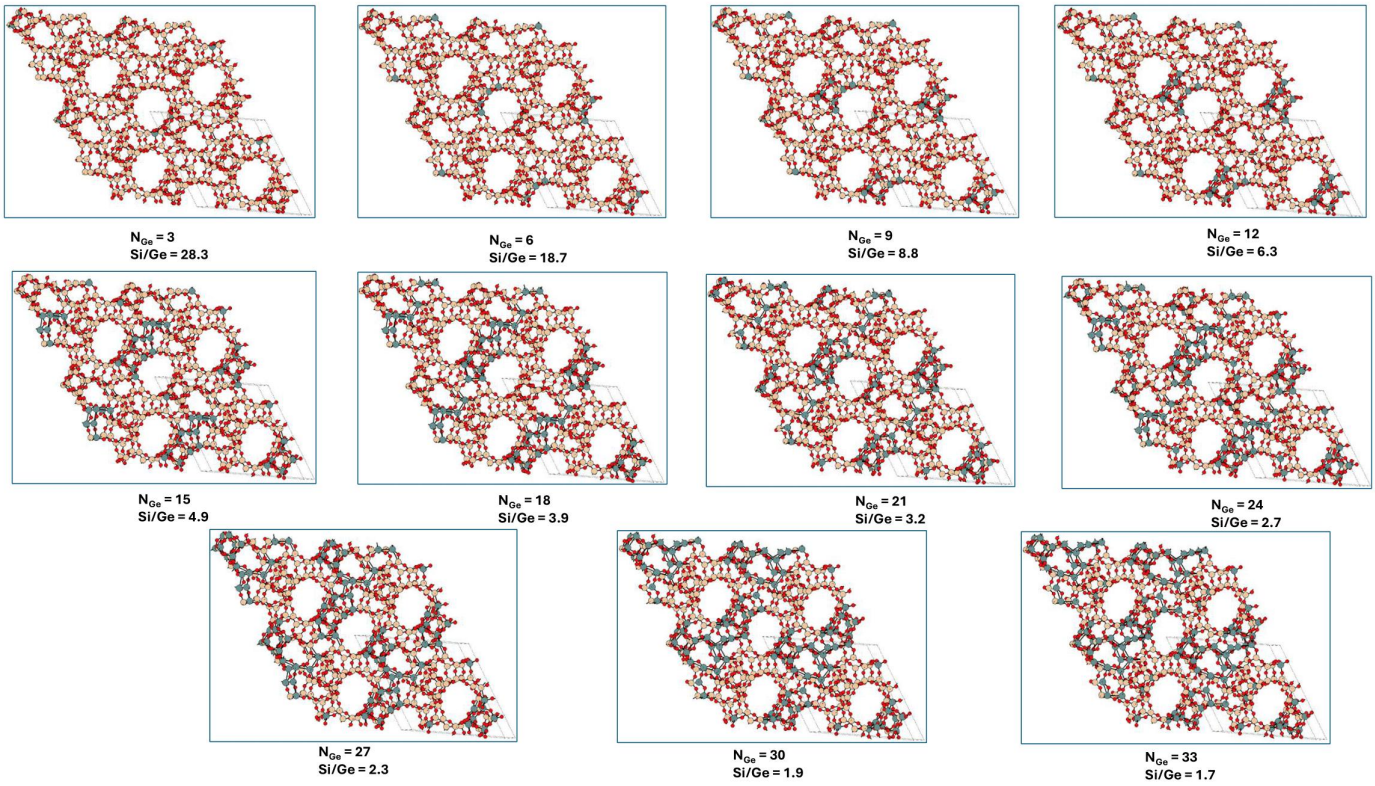


Fig. S21 The global structure optima for UOV zeolite (single-cell model) at all Si/Ge ratios considered in this work, where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

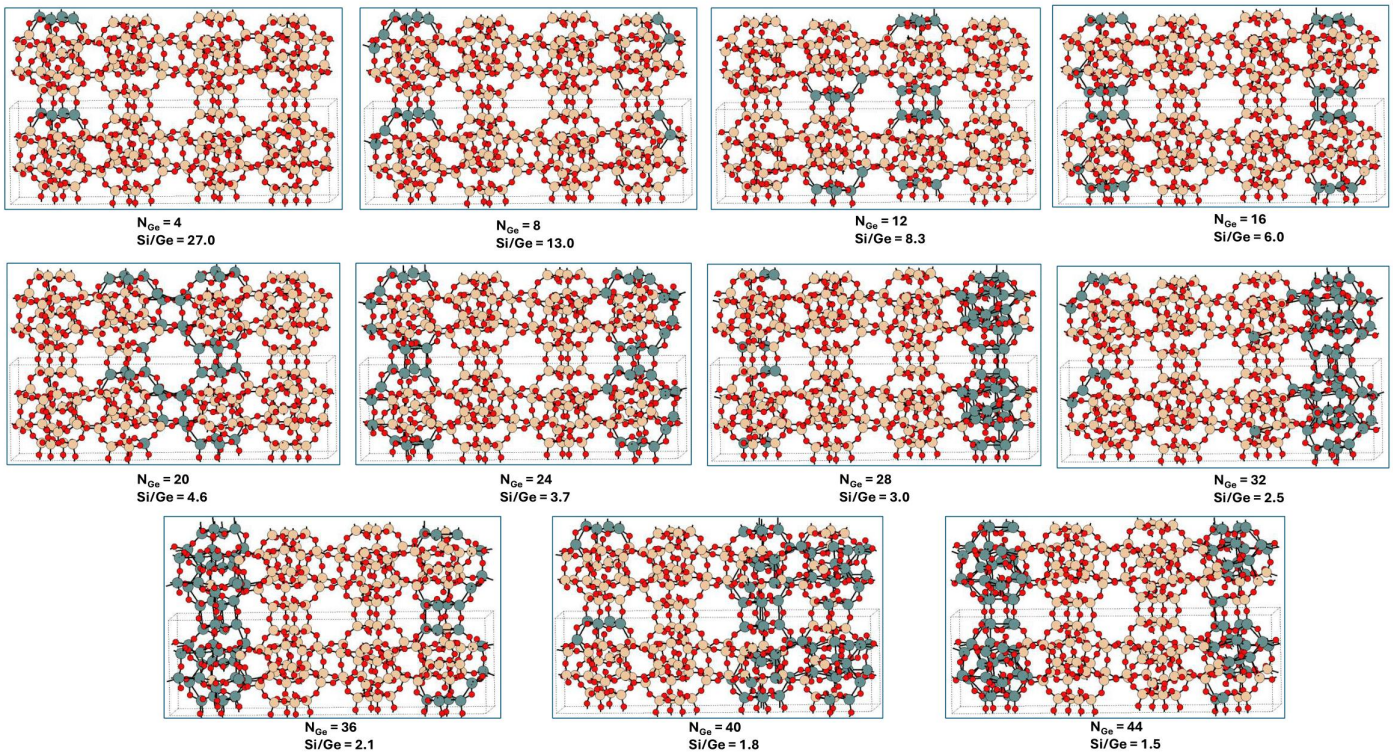


Fig. S22 The global structure optima for IWW zeolite (single-cell model) at all Si/Ge ratios considered in this work, where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the b-axis is horizontal, and the c-axis is vertical.

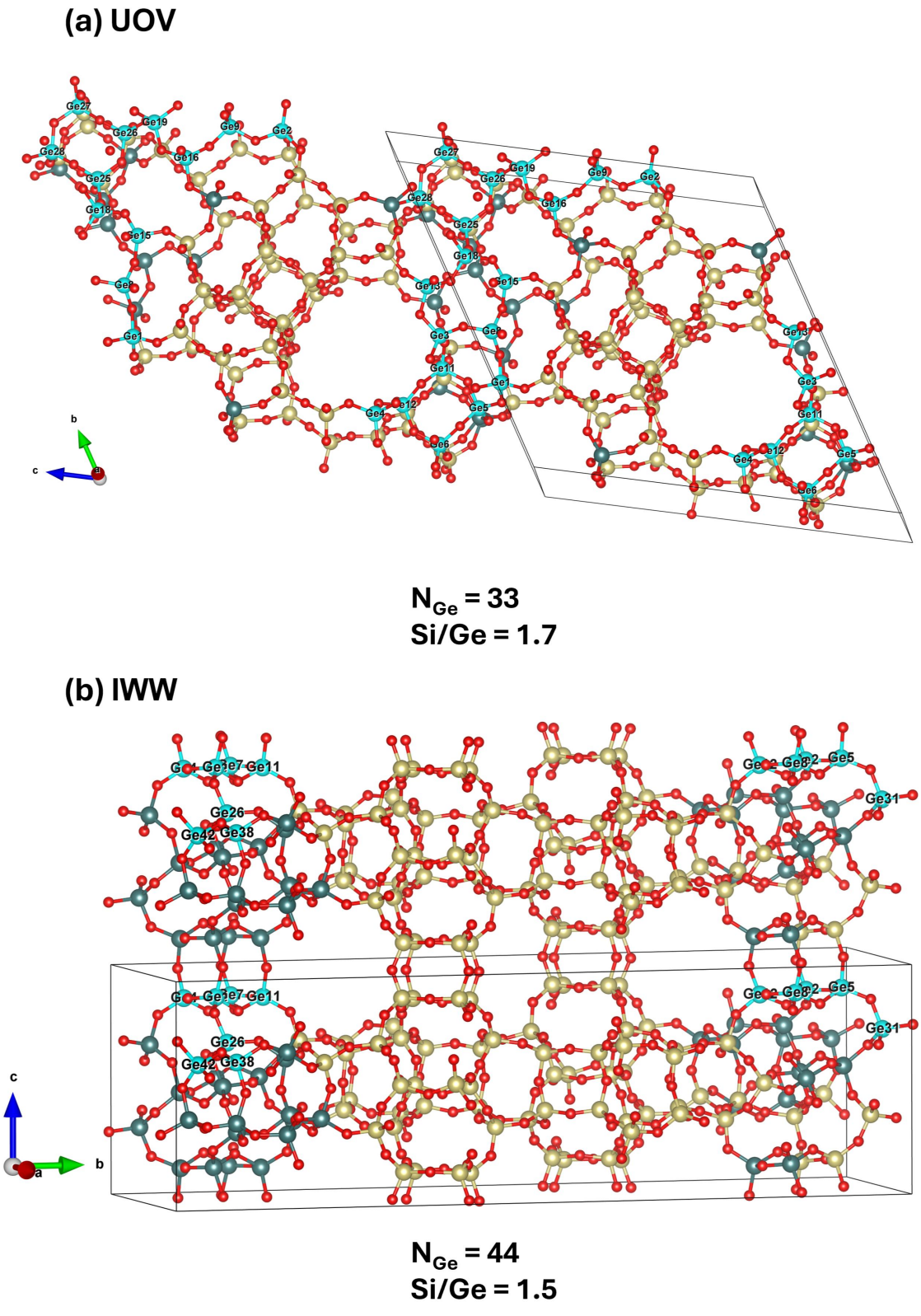


Fig. S23 Low energy structures (with relative energy $E_{rel} < 10$ kJ/mol) for UOV and IWW zeolites (a single-cell model, where N_{Ge}, No. of Ge/Unit Cell) highlighting (numbered) the Single-Four-Membered Rings (S4Rs) occupied heavily with germanium atoms (cyan).

5.1 UTL

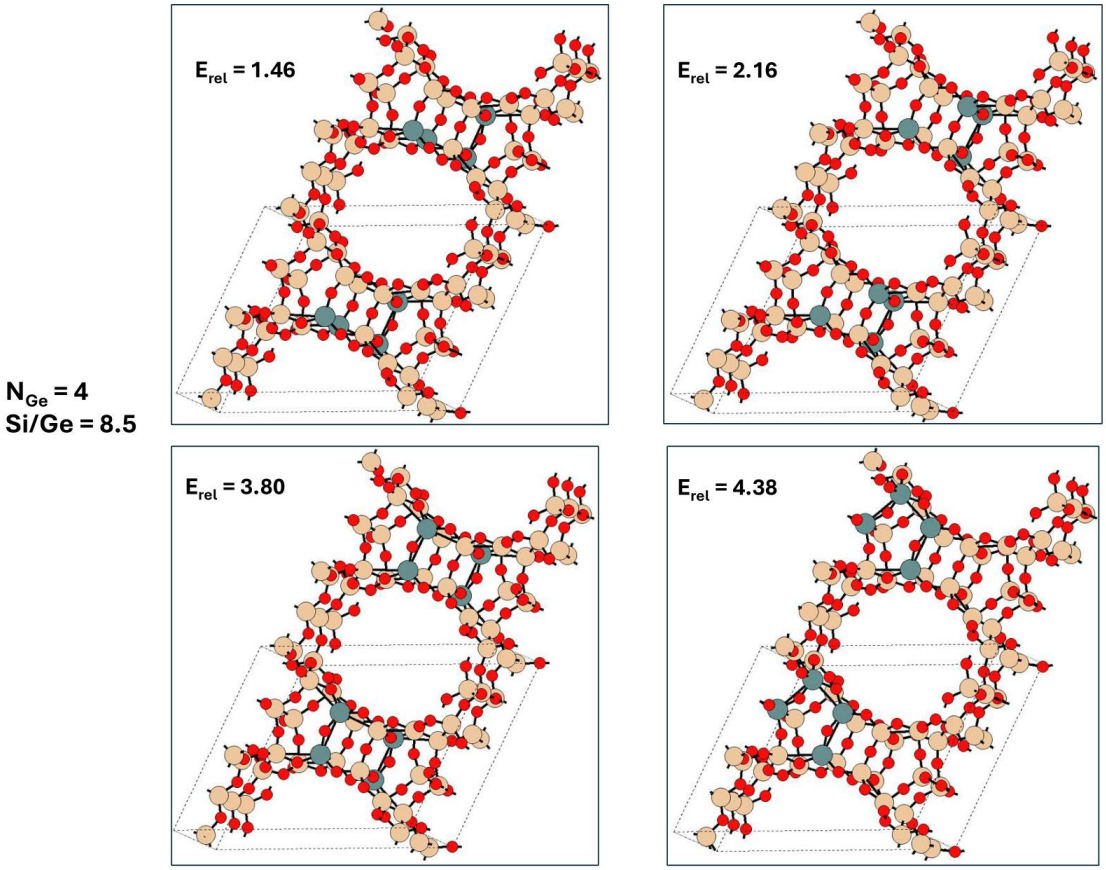


Fig. S24 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for UTL zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to half-filling of all D4R unit T-sites. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

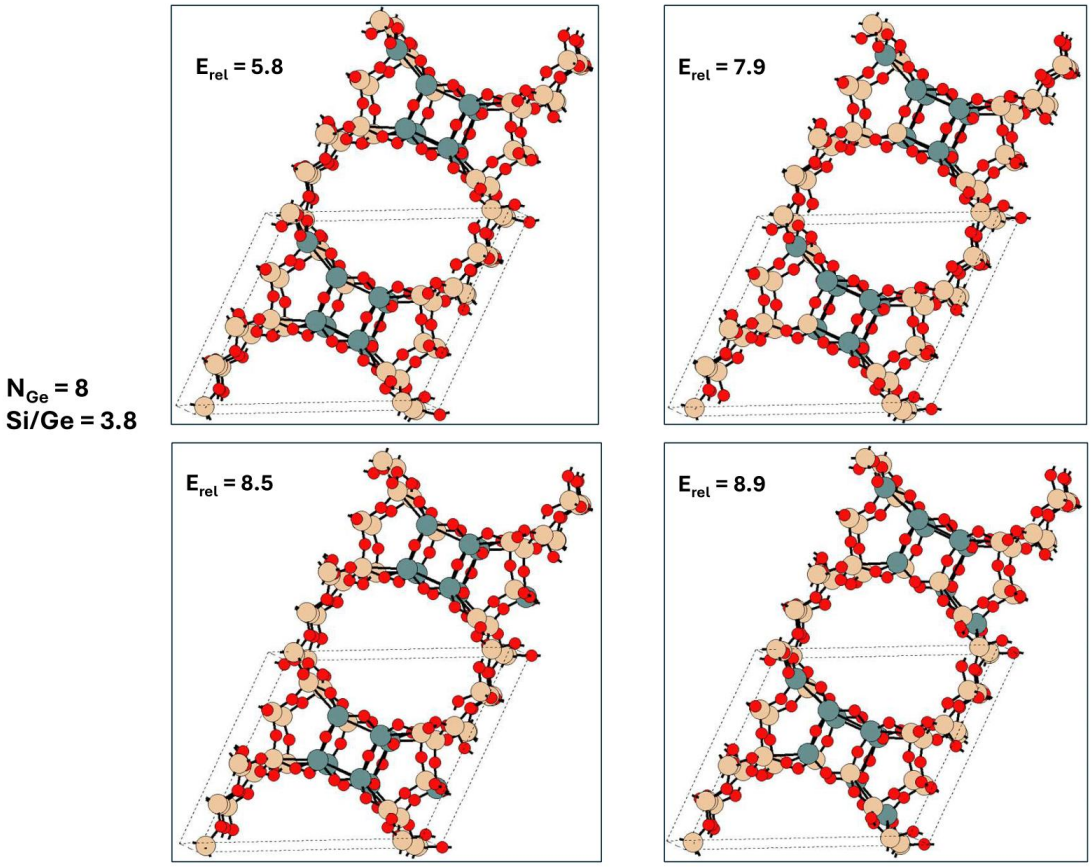


Fig. S25 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for UTL zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to full-filling of all D4R unit T-sites. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

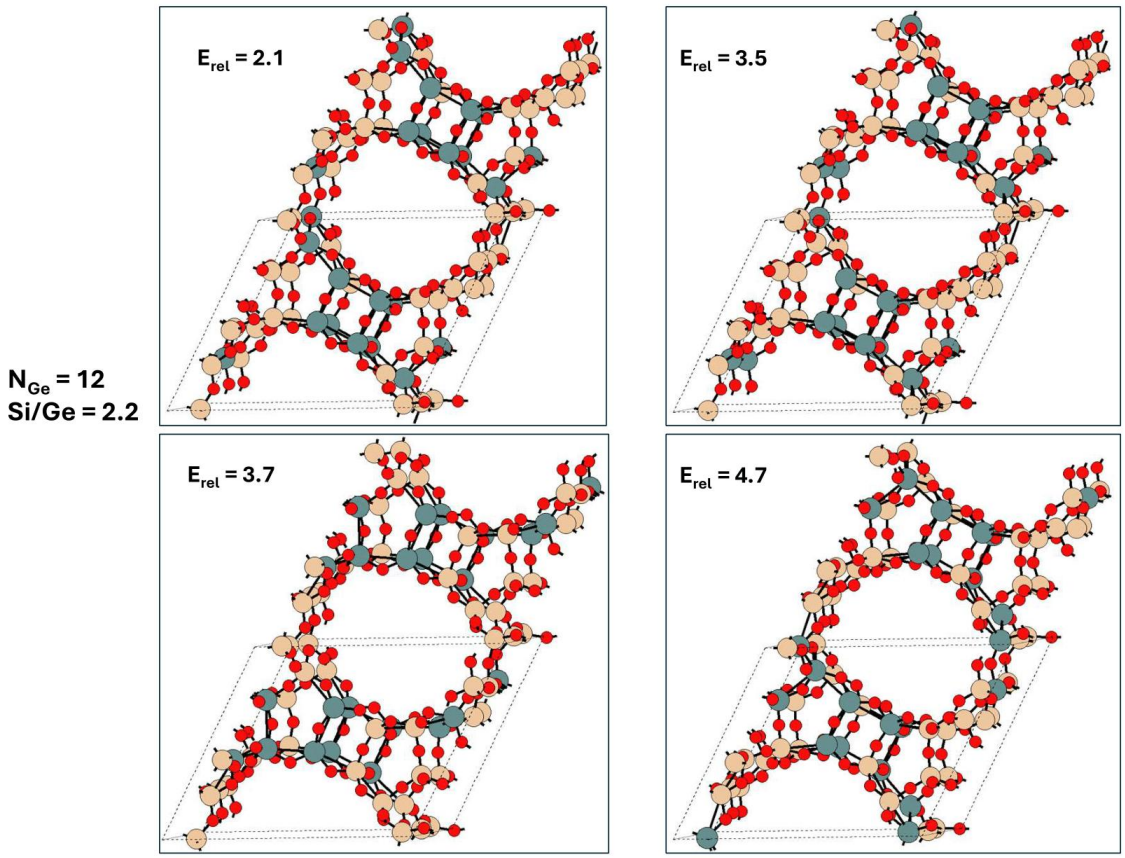


Fig. S26 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for UTL zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is maximum in our simulation. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

5.2 BEC

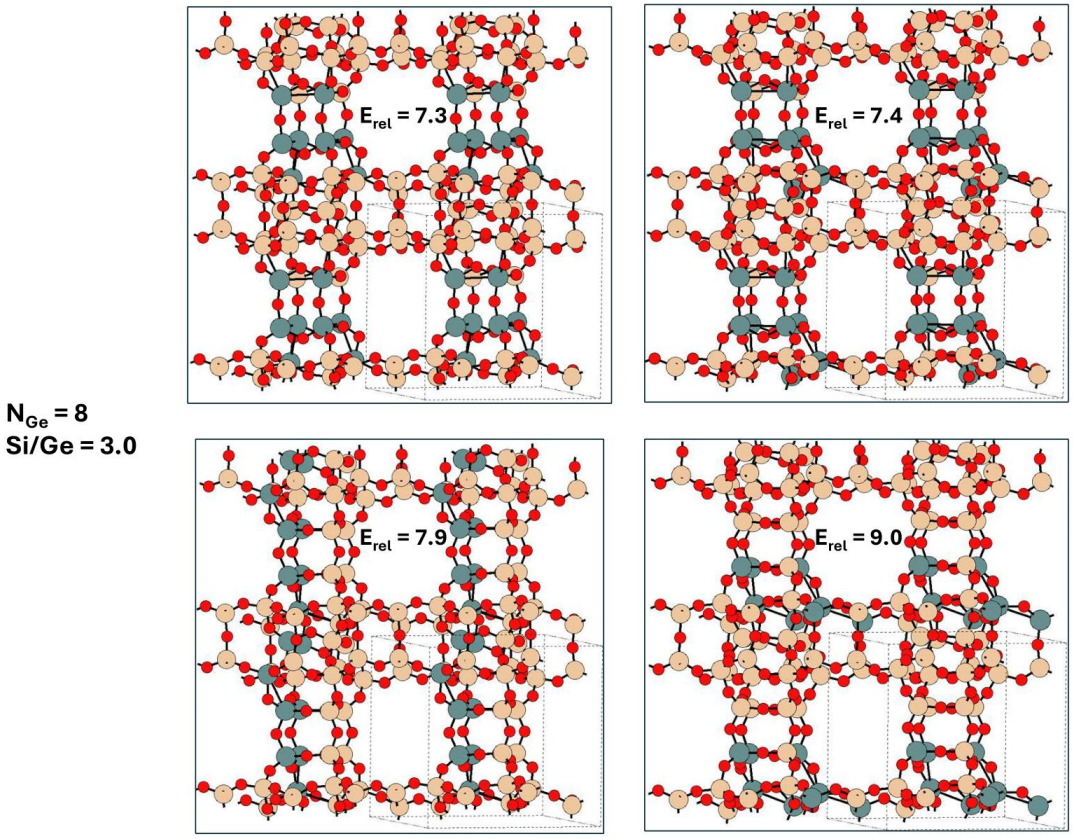


Fig. S27 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for BEC zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to half-filling of all D4R unit T-sites. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

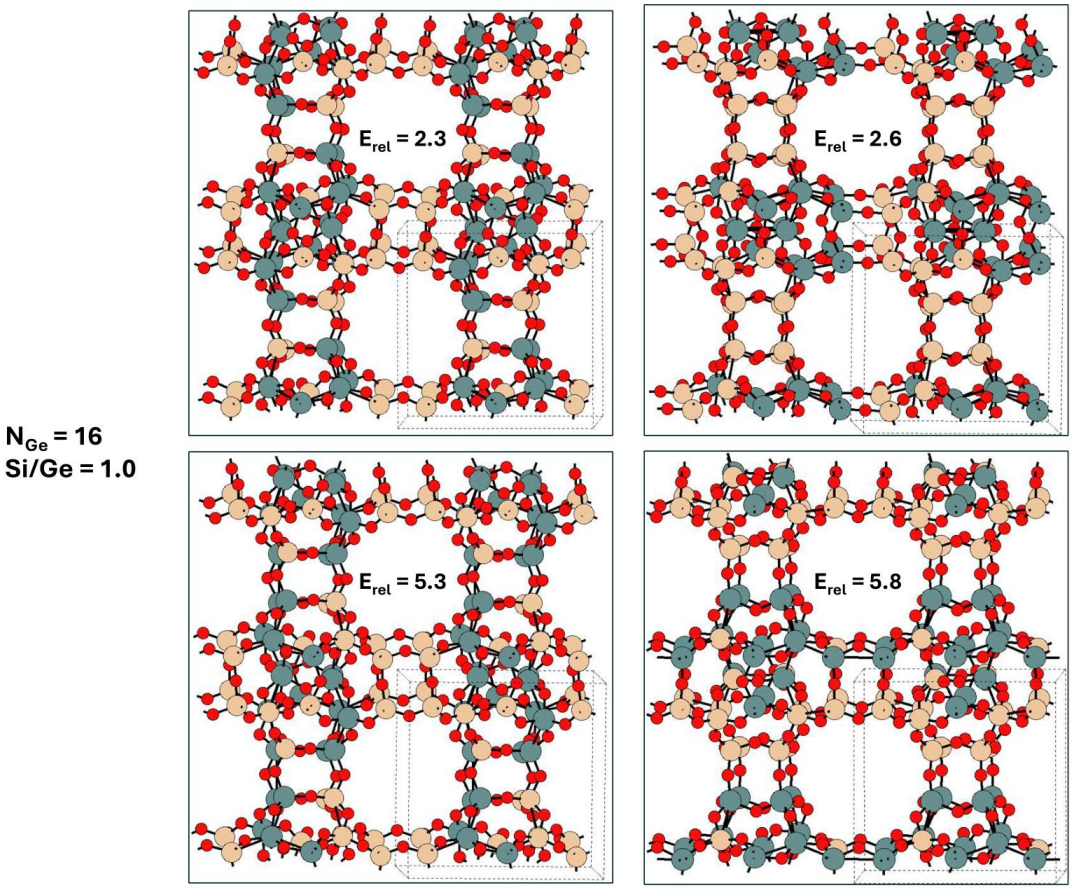


Fig. S28 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for BEC zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to full-filling of all D4R unit T-sites, in this case it is also the maximum germanium loading considered in our simulation. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

5.3 CTH

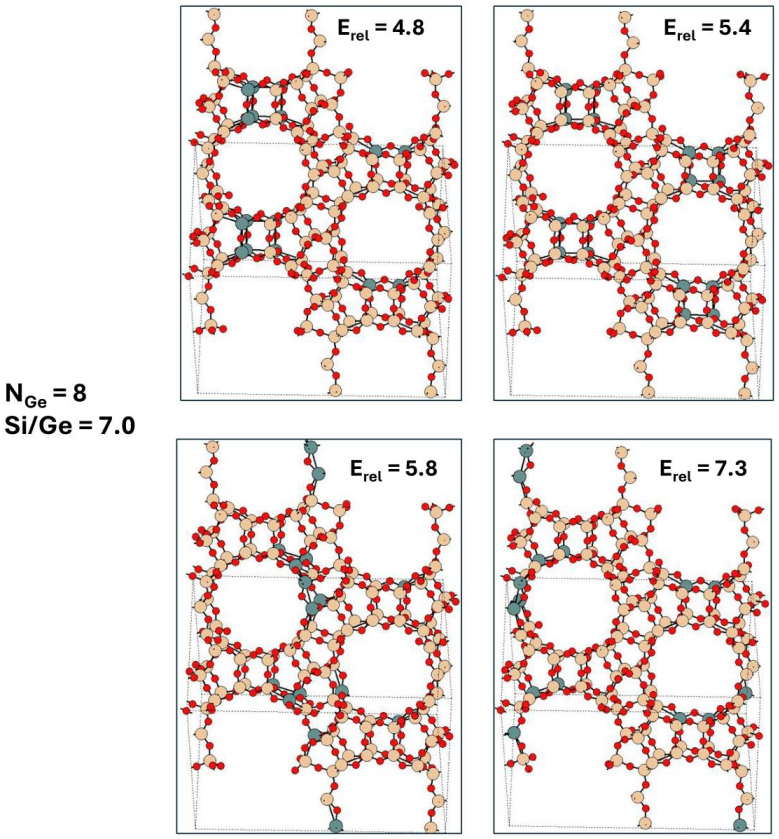


Fig. S29 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for CTH(-A) zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to half-filling of all D4R unit T-sites. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

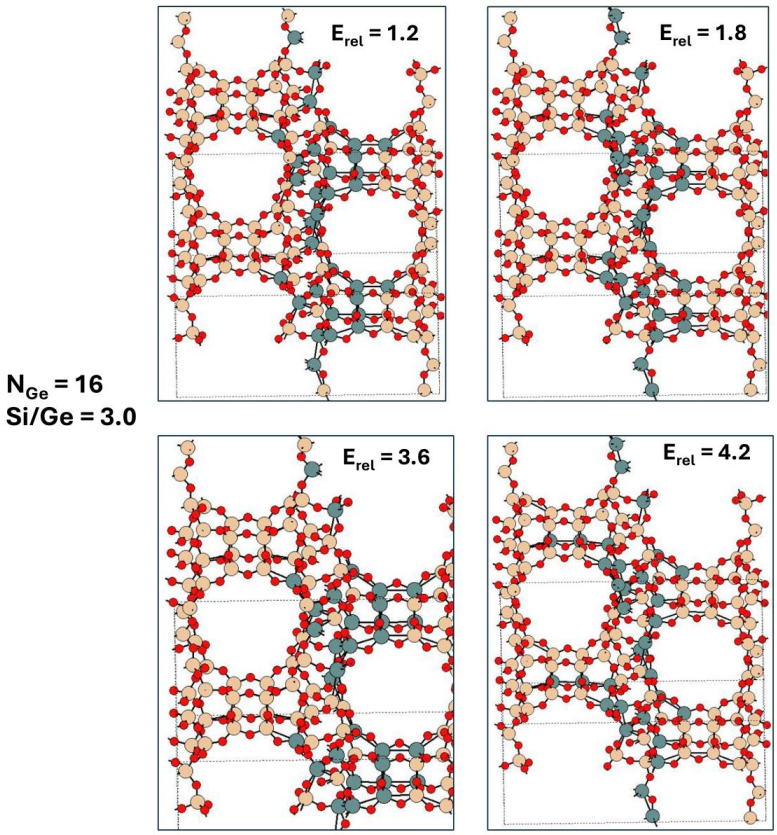


Fig. S30 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for CTH(-A) zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to full-filling of all D4R unit T-sites. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

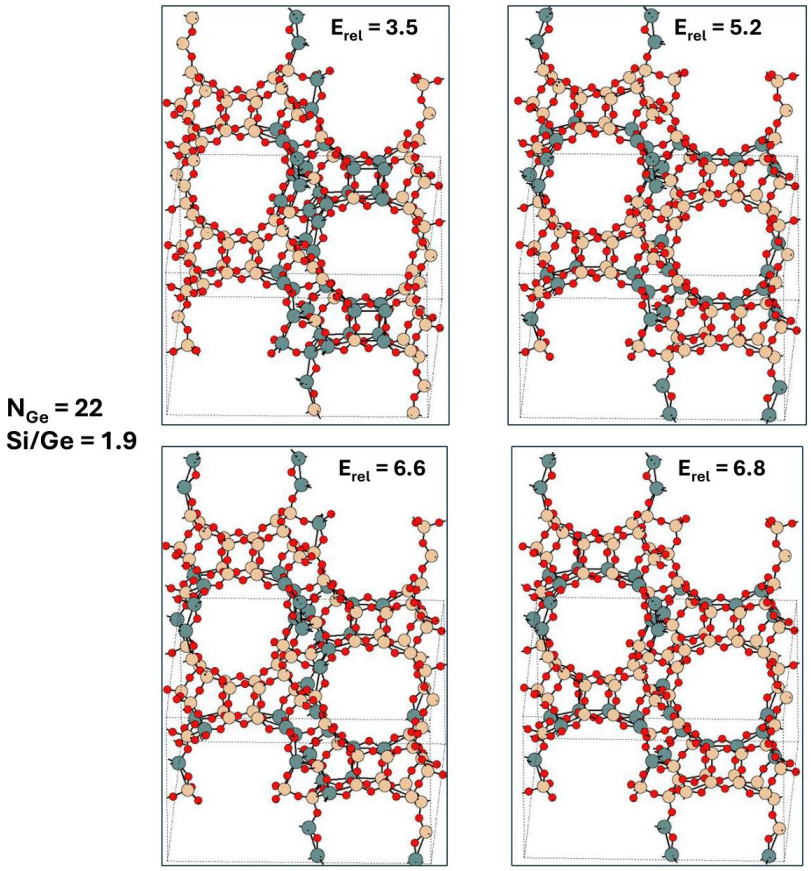


Fig. S31 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for CTH(-A) zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to the maximum germanium considered in our simulation. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

5.4 UOV

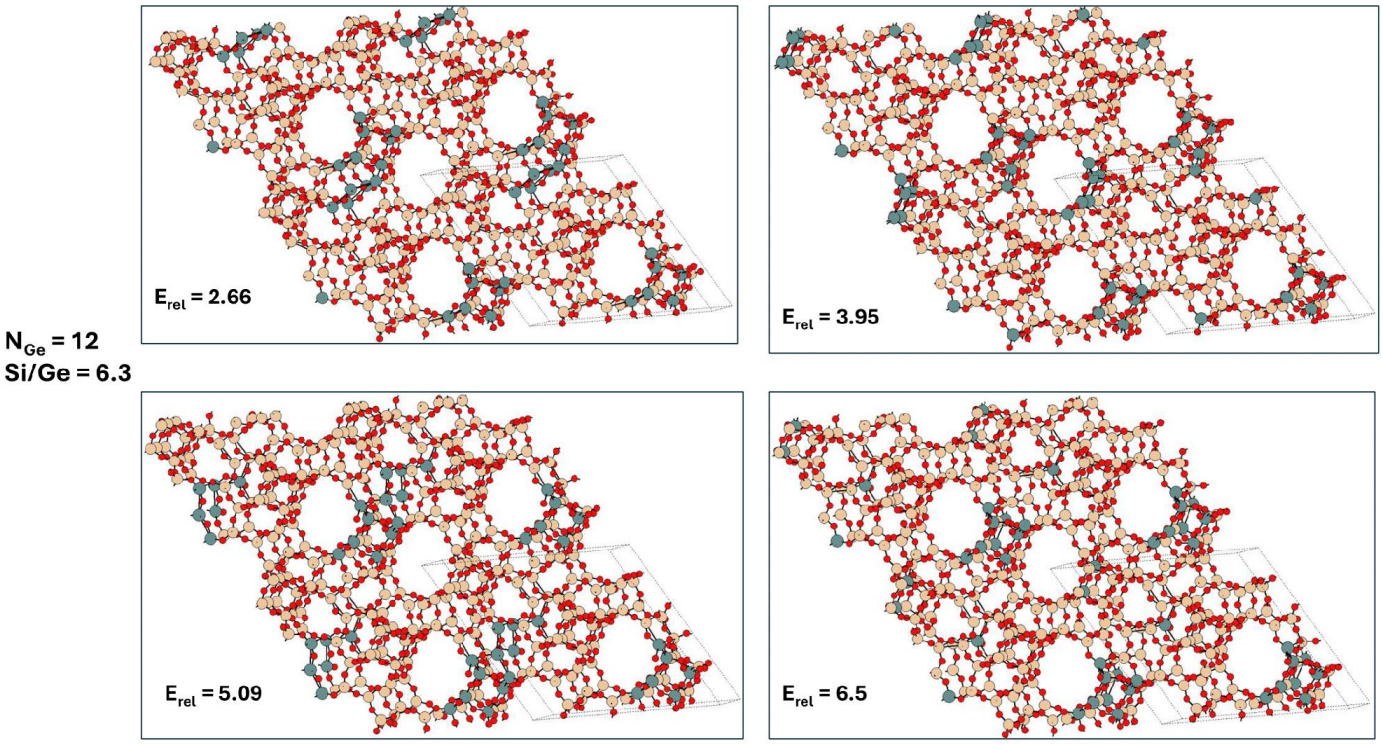


Fig. S32 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for UOV zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to half-filling of all D4R unit T-sites. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

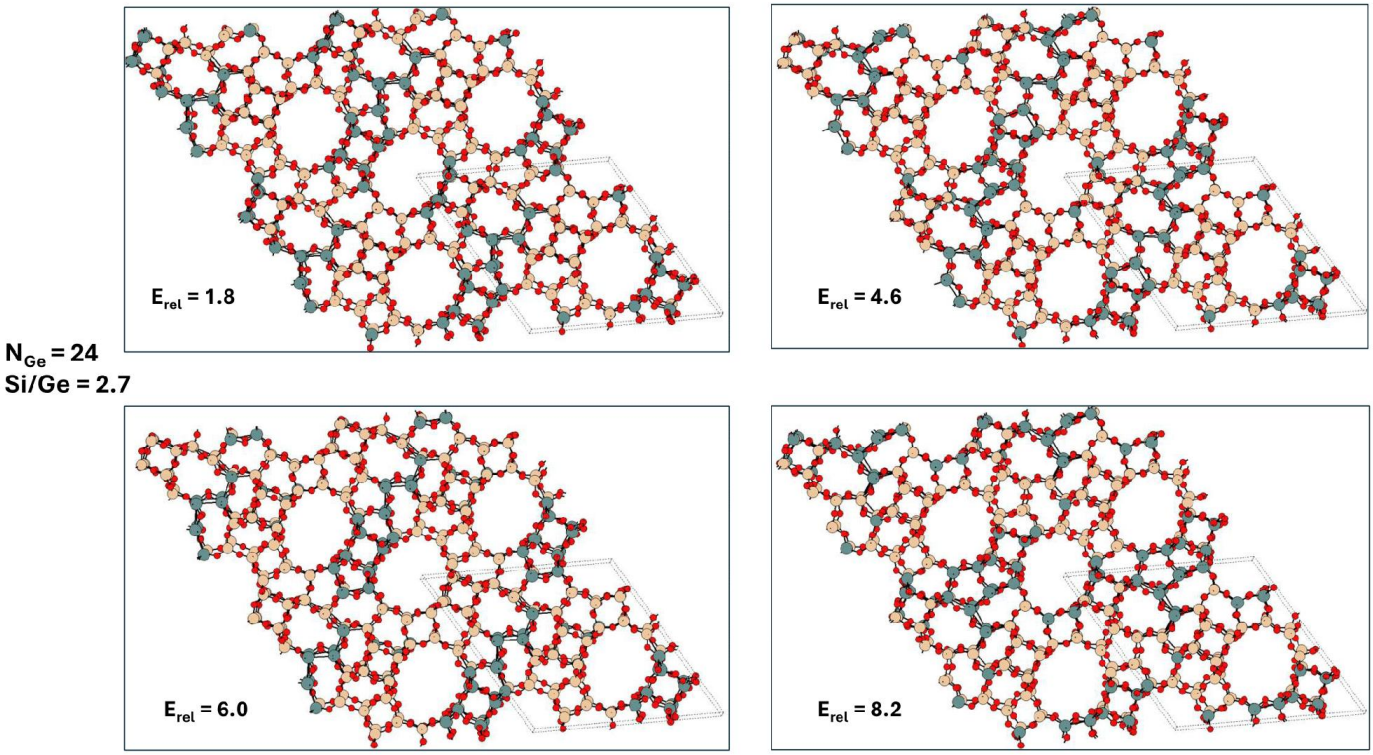


Fig. S33 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for UOV zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to full-filling of all D4R unit T-sites. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

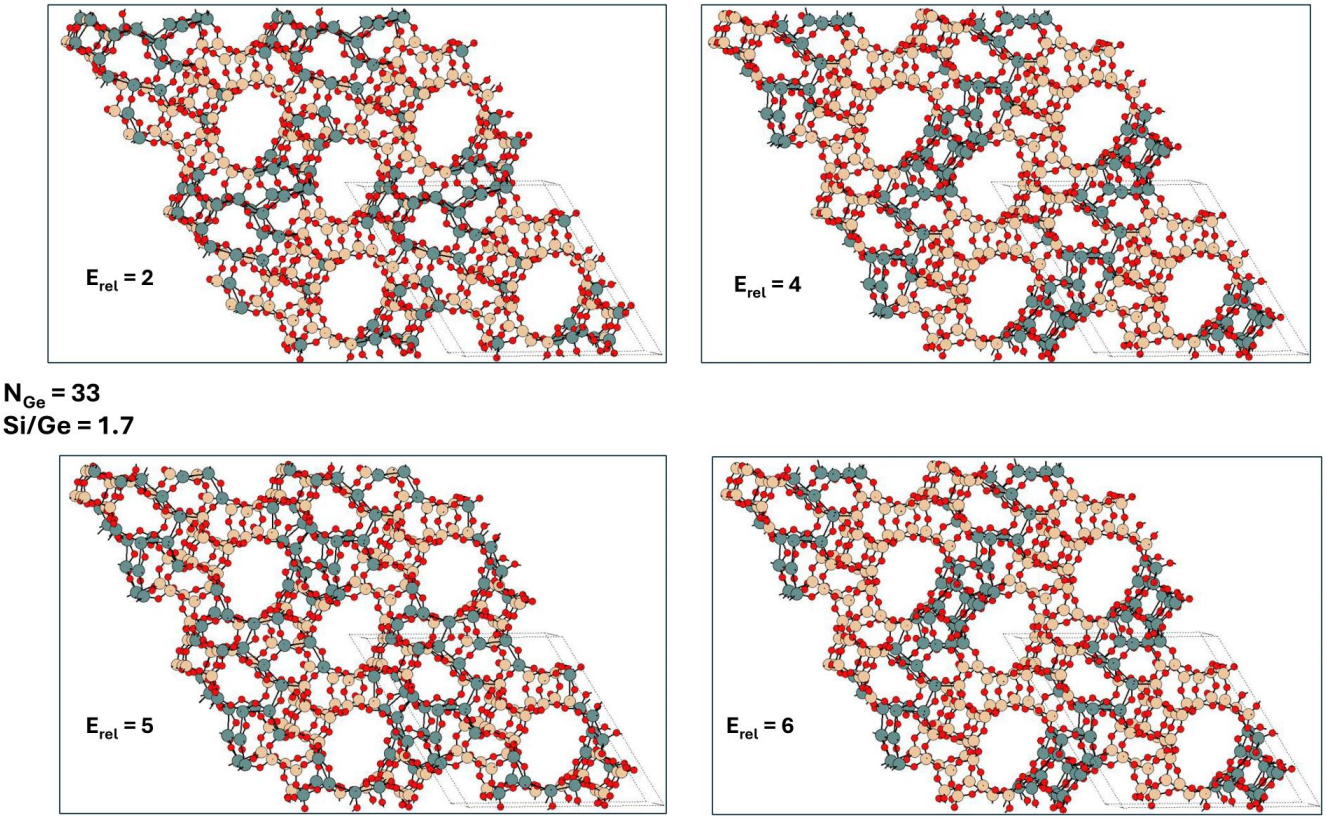


Fig. S34 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for UOV zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to the maximum germanium considered in our simulation. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

5.5 IWW

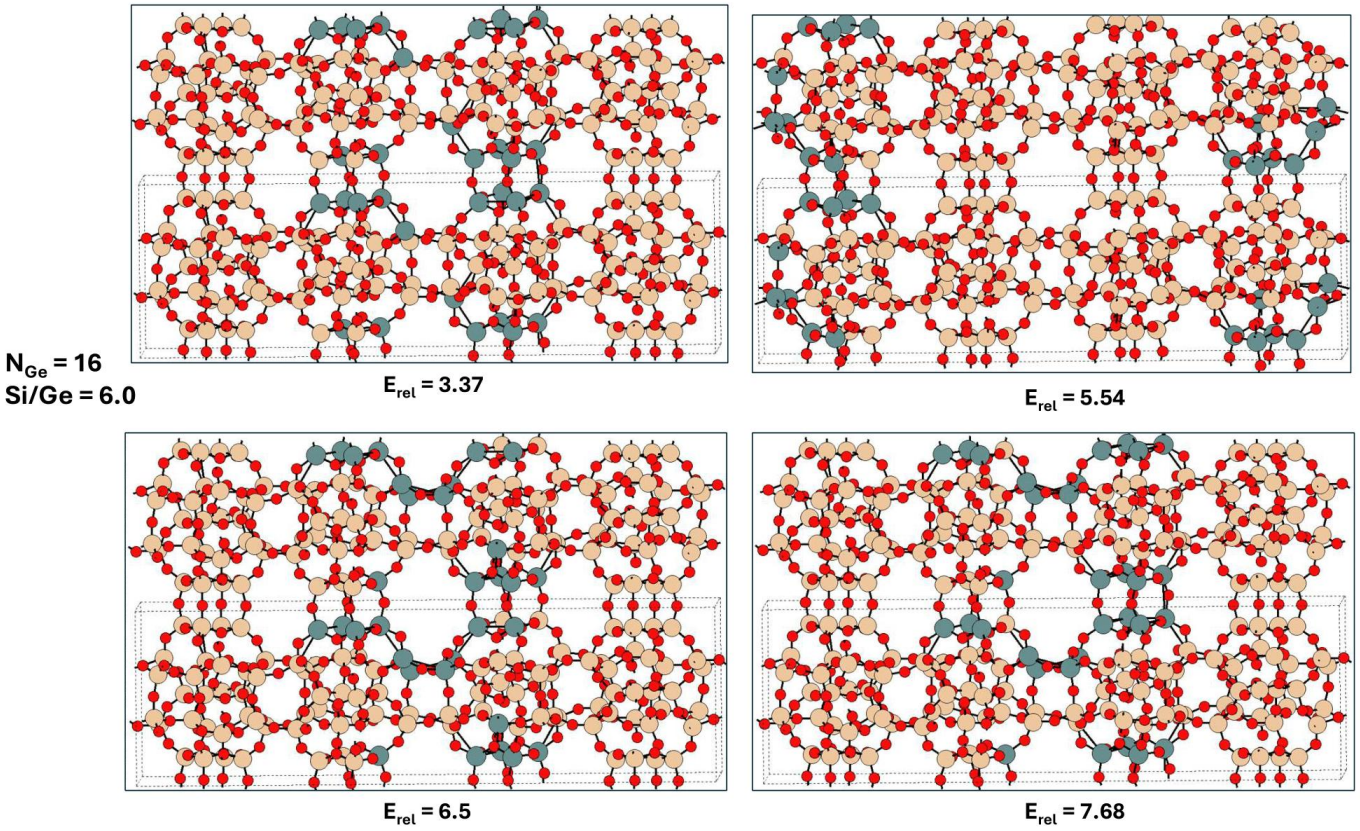


Fig. S35 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for IWW zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to half-filling of all D4R unit T-sites. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the b-axis is horizontal, and the c-axis is vertical.

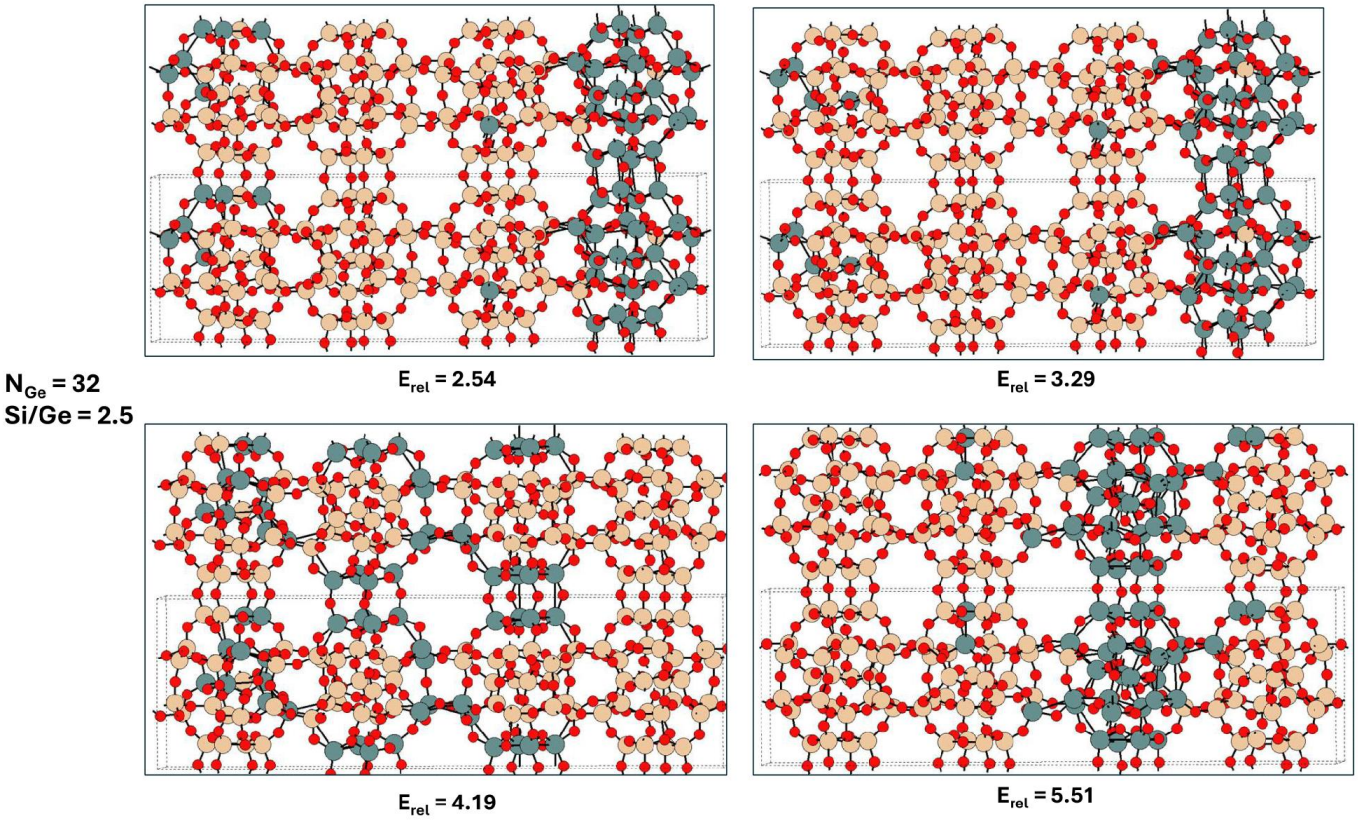


Fig. S36 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for IWW zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to full-filling of all D4R unit T-sites. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the b-axis is horizontal, and the c-axis is vertical.

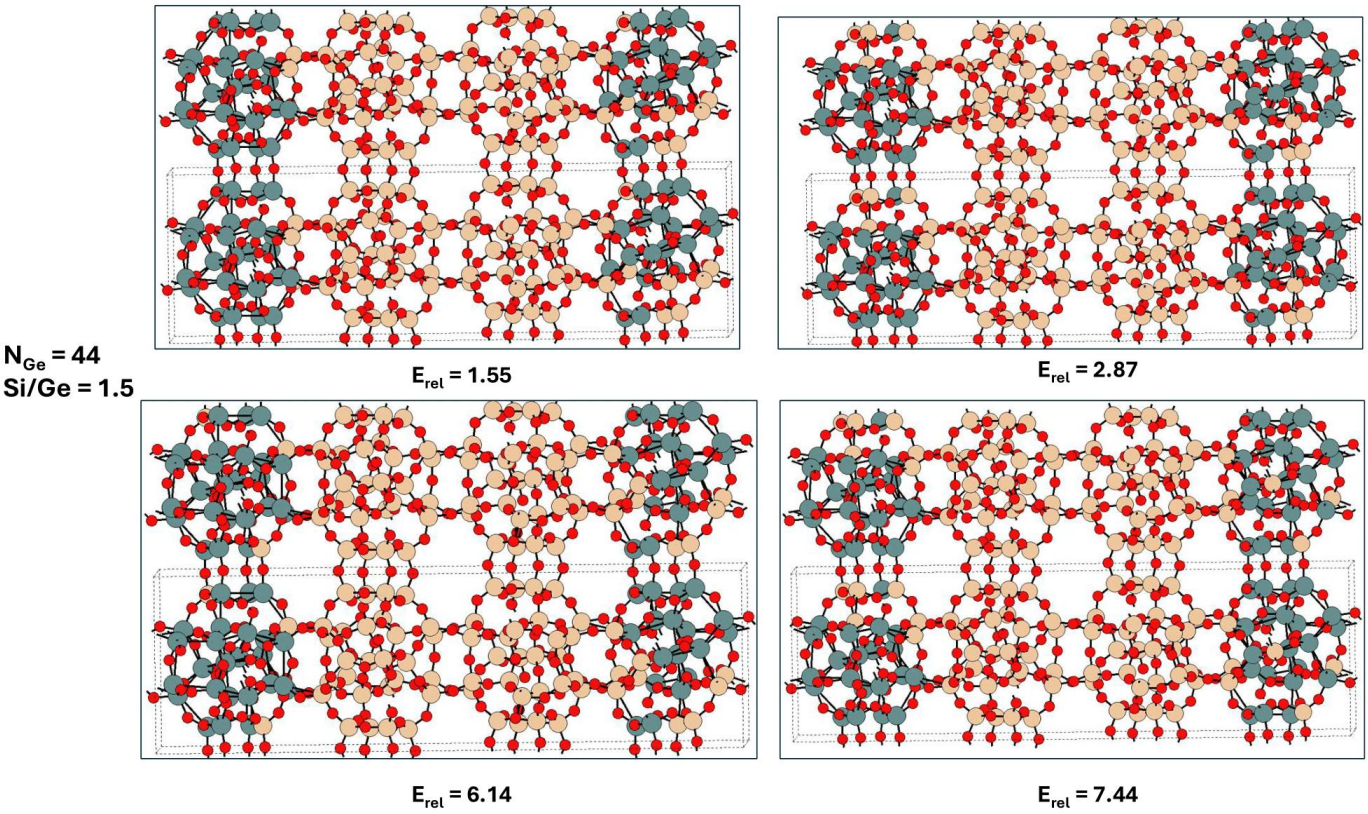


Fig. S37 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for IWW zeolite (single-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to the maximum germanium considered in our simulation. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the b-axis is horizontal, and the c-axis is vertical.

6 Results: Finite Size Effects (Supercell)

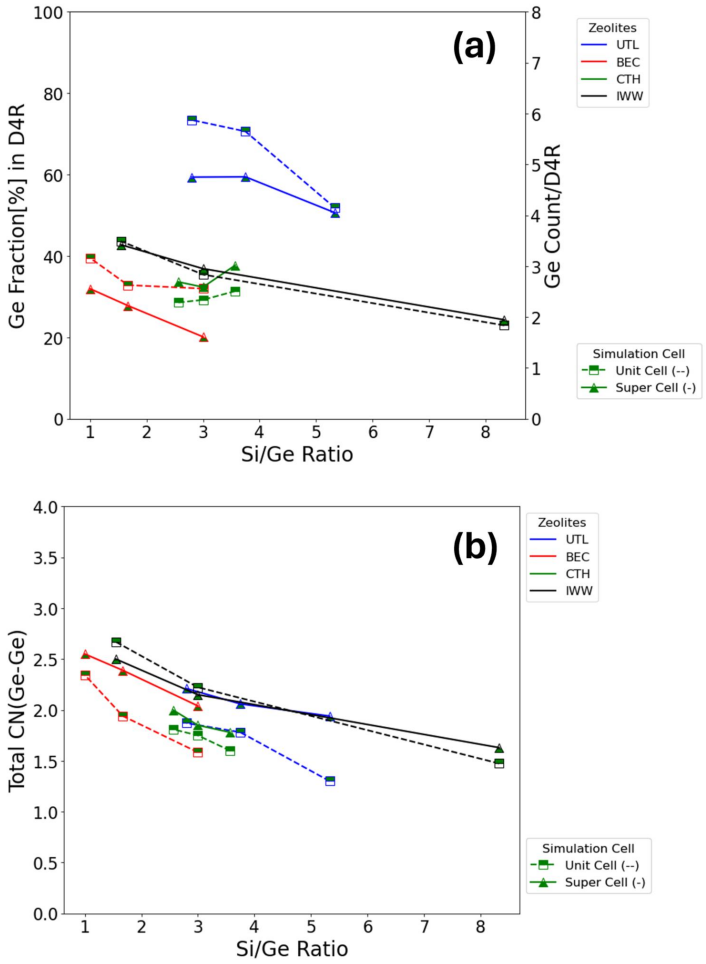


Fig. S38 Quantification of the D4R occupation and the clustering tendency (irrespective of the T-site type) of germanium as a function of the simulation cell size for UTL, BEC, CTH and IWW zeolite topologies for a few selected values of Si/Ge ratio, using the metrics defined in Section 2.3, namely: (a) Ge Fraction[%] in D4R and Ge Count/D4R in primary and secondary y-axis, respectively and (b) Total CN(Ge-Ge): Ge-Ge coordination number

.

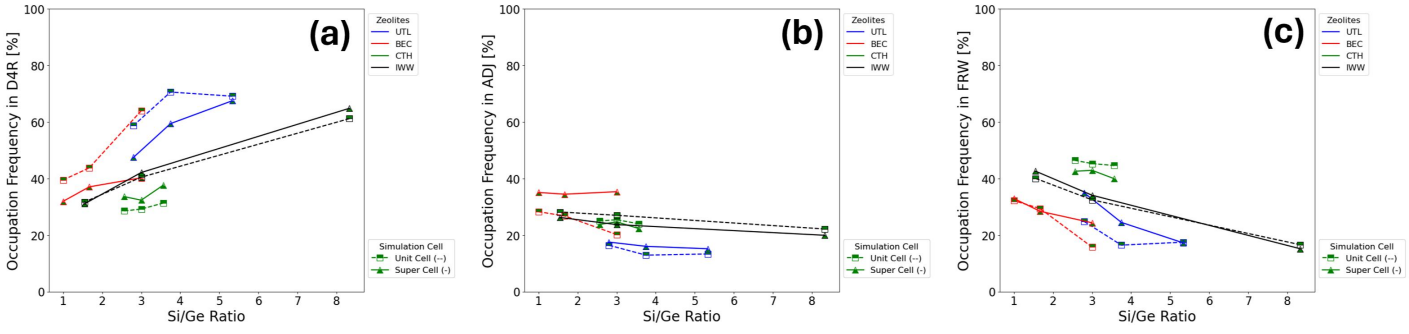


Fig. S39 Quantification of the finite size effects on one of the characteristics of the germanium distribution, the occupation frequency [in %] (see Section 2.3 for definition) of germanium atoms in different types of T sites - (a) *D4R*, (b) *adjacent*, (c) *framework* (see also Figure SI S1)

6.1 UTL Supercell

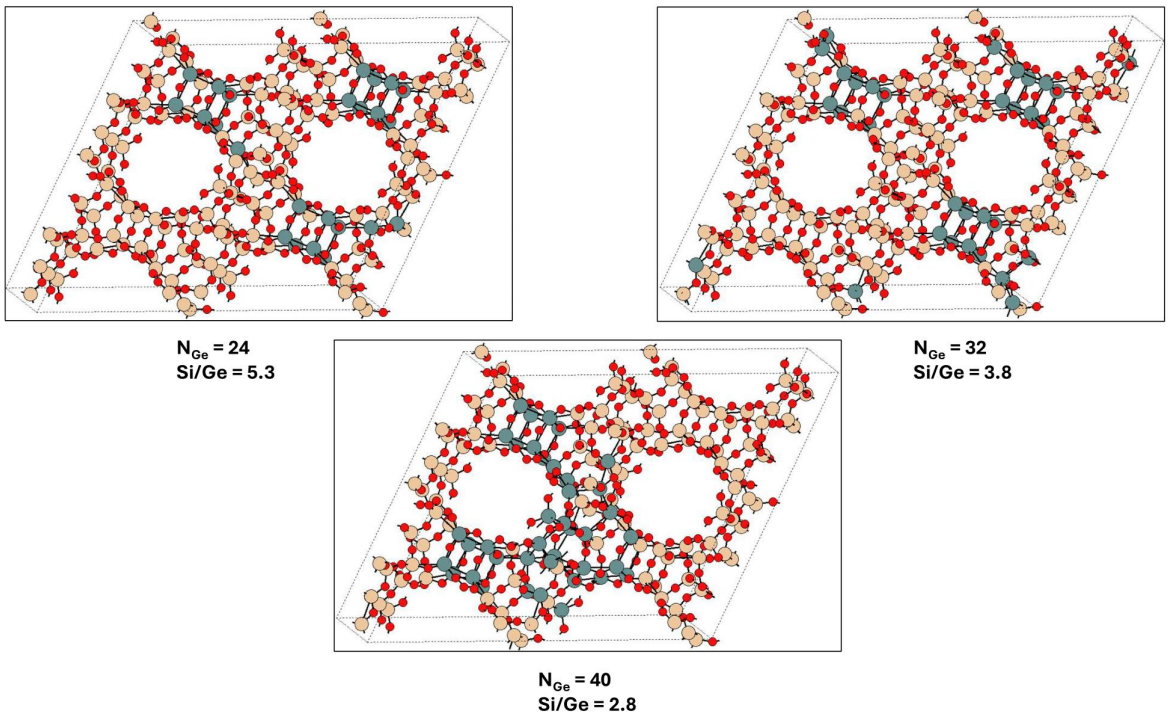


Fig. S40 The GSO structures for UTL zeolite (super-cell model) at various Si/Ge where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure: viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

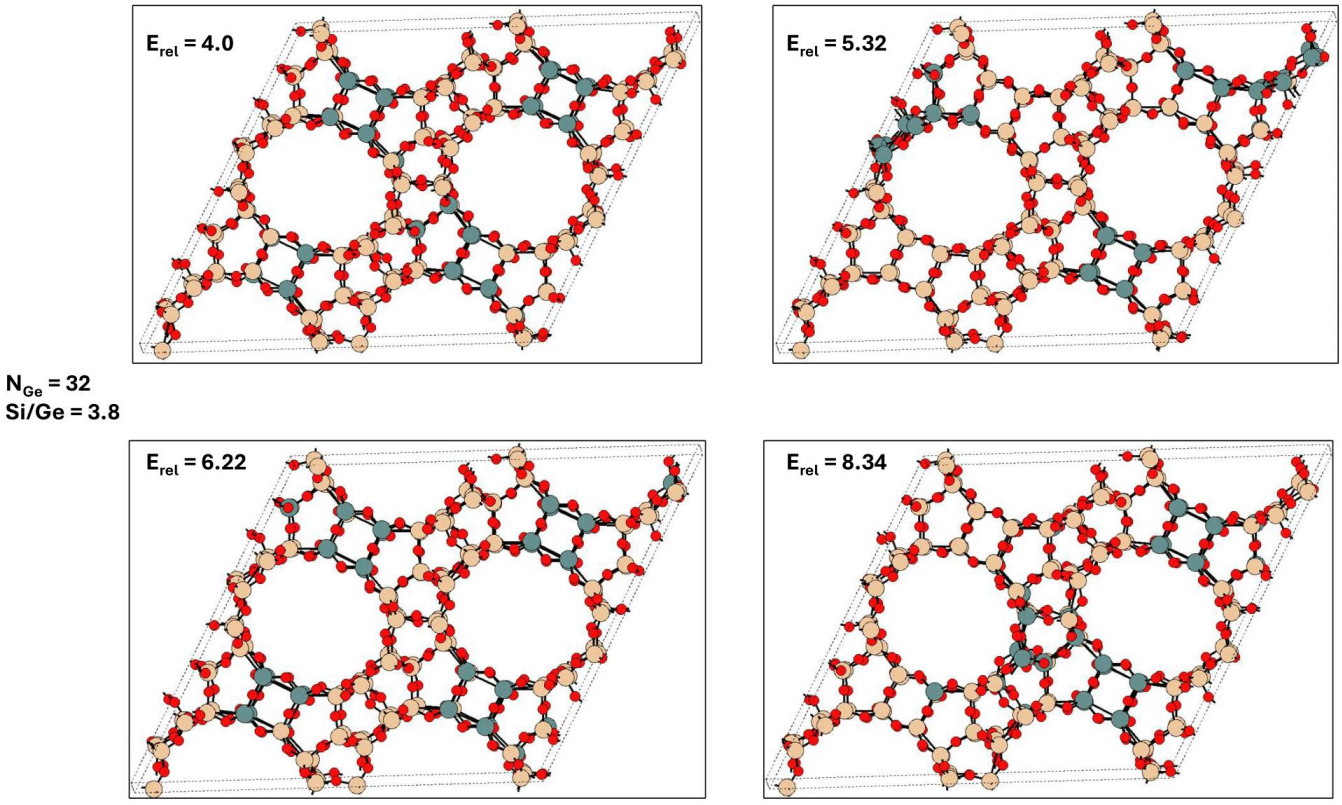


Fig. S41 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for UTL zeolite (super-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to the full-filling of all D4R unit T-sites. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

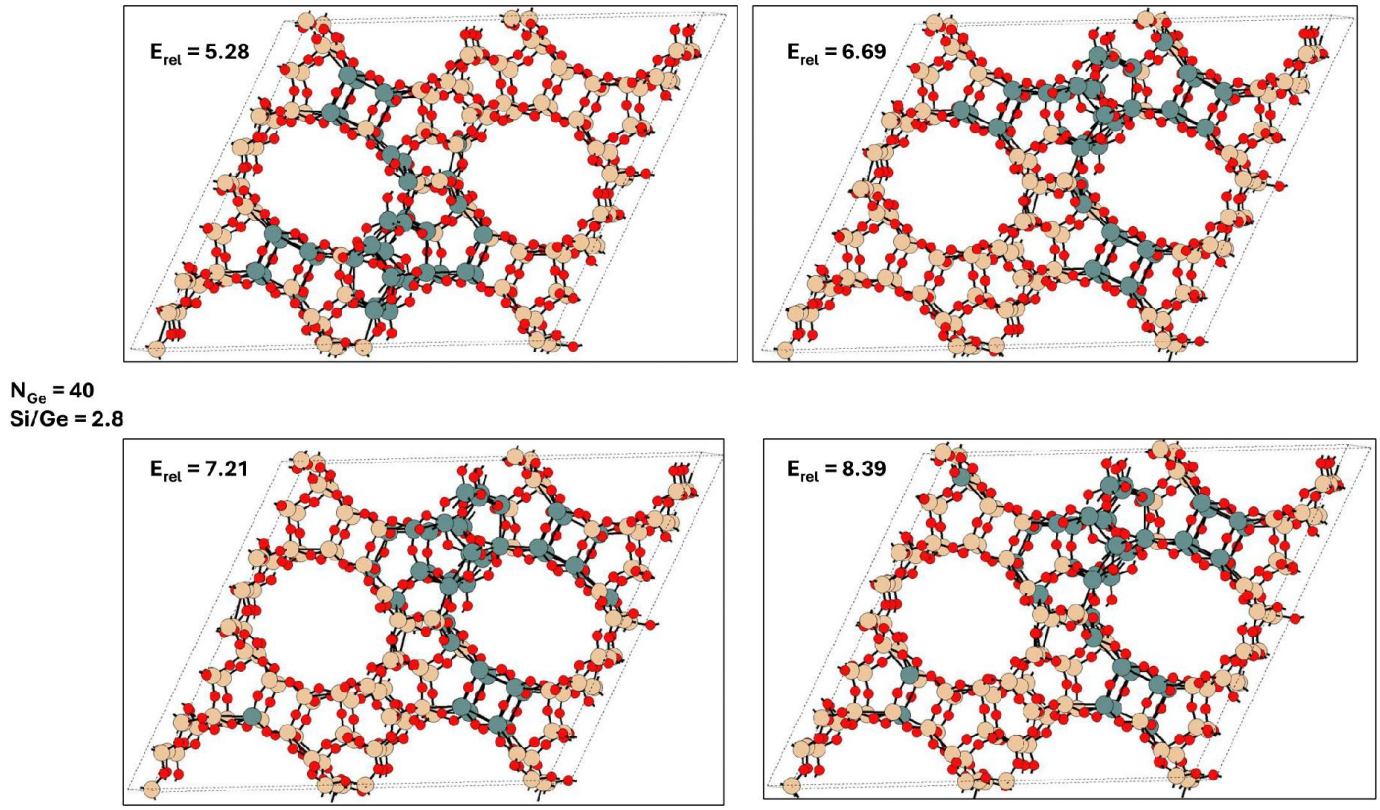


Fig. S42 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for UTL zeolite (super-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to maximum germanium considered in our simulation. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

6.2 BEC Supercell

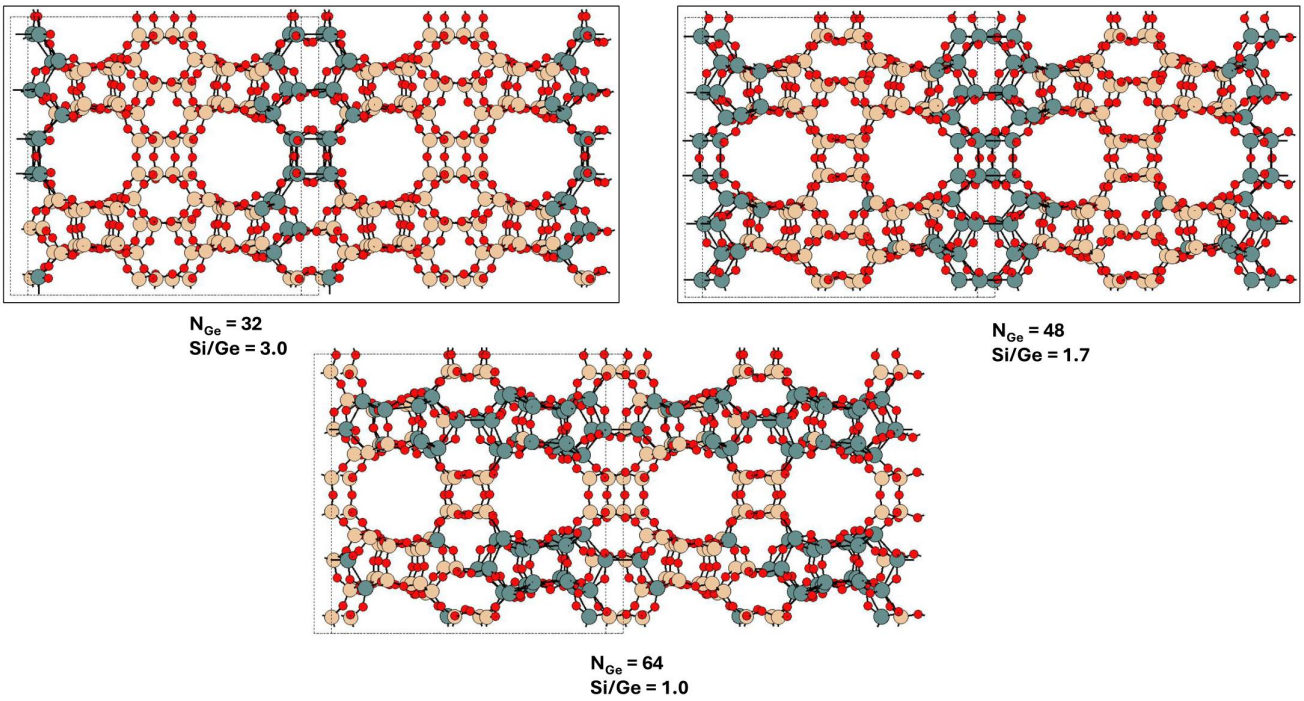


Fig. S43 The GSO structures for BEC zeolite (super-cell model) at various Si/Ge where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

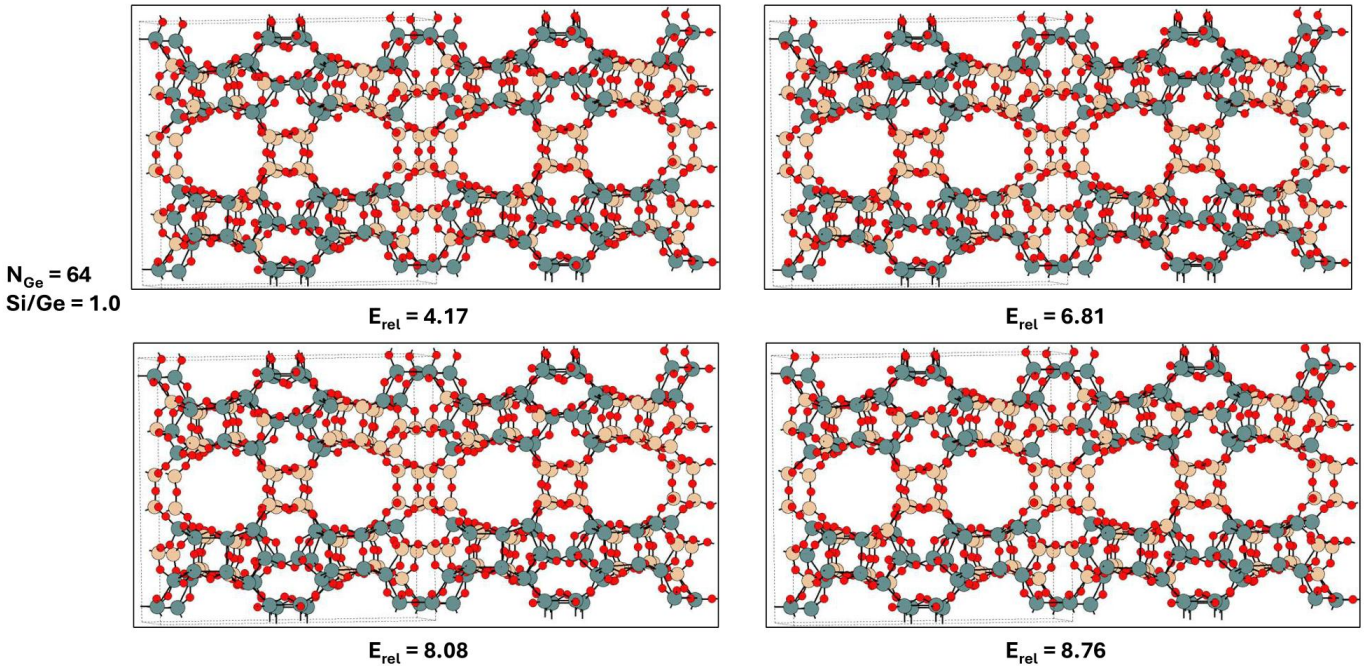


Fig. S44 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for BEC zeolite (super-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to the full-filling of all D4R unit T-sites, which is also the maximum germanium loading considered in our simulation. The zeolite structure is viewed along the c-axis, onto the a-b plane. In this orientation, the a-axis is horizontal, and the b-axis is vertical.

6.3 CTH Supercell

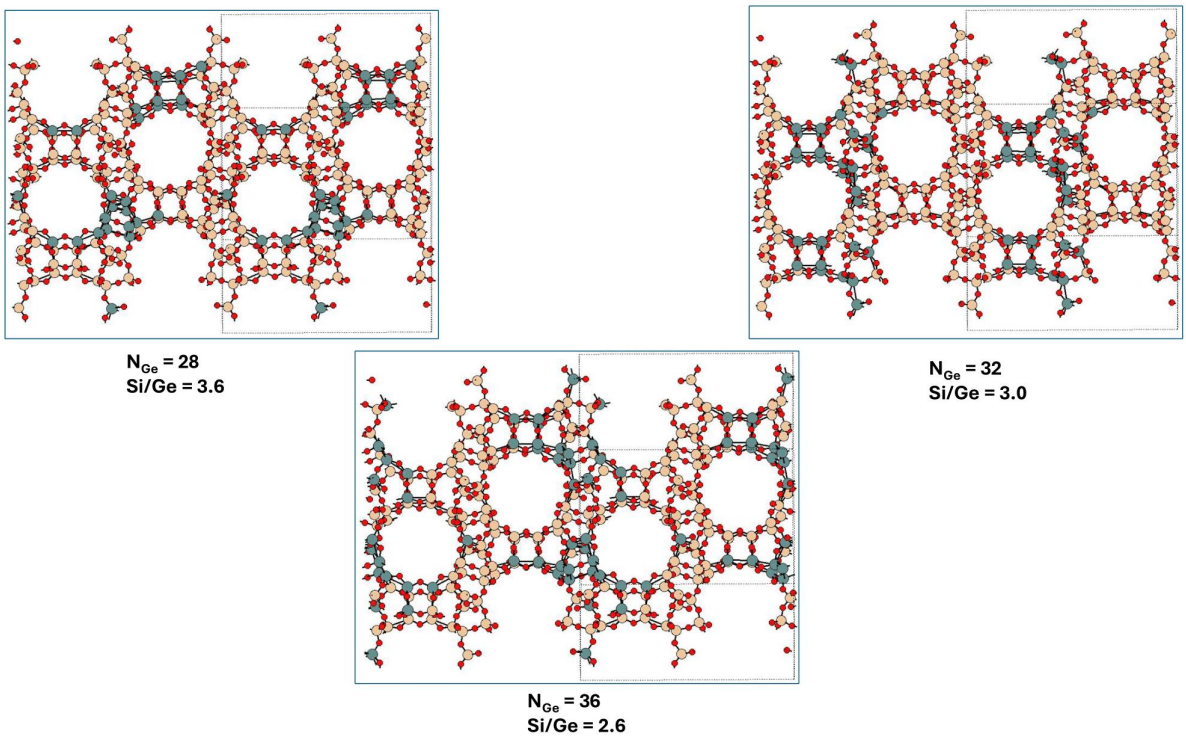


Fig. S45 The GSO structures for CTH(-A) zeolite (super-cell model) at various Si/Ge where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

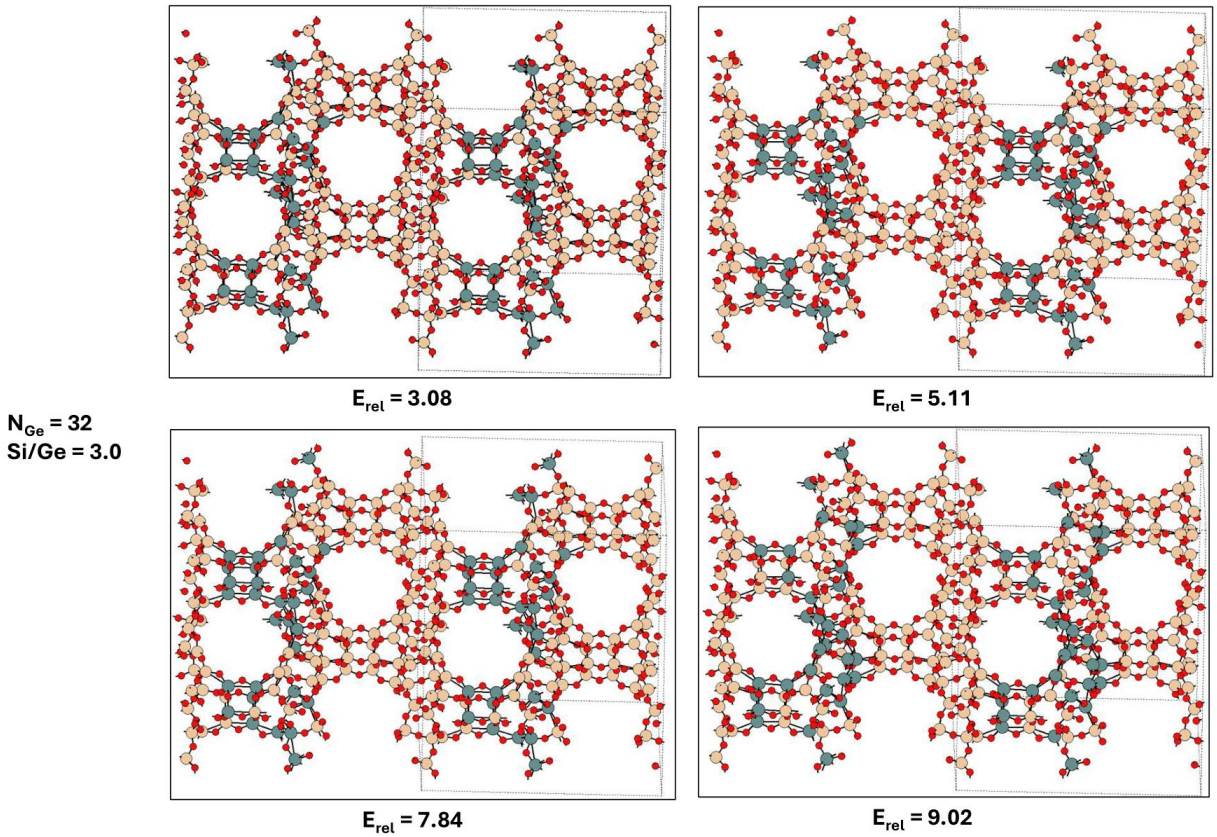


Fig. S46 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for CTH(-A) zeolite (super-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to the full-filling of all D4R unit T-sites. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

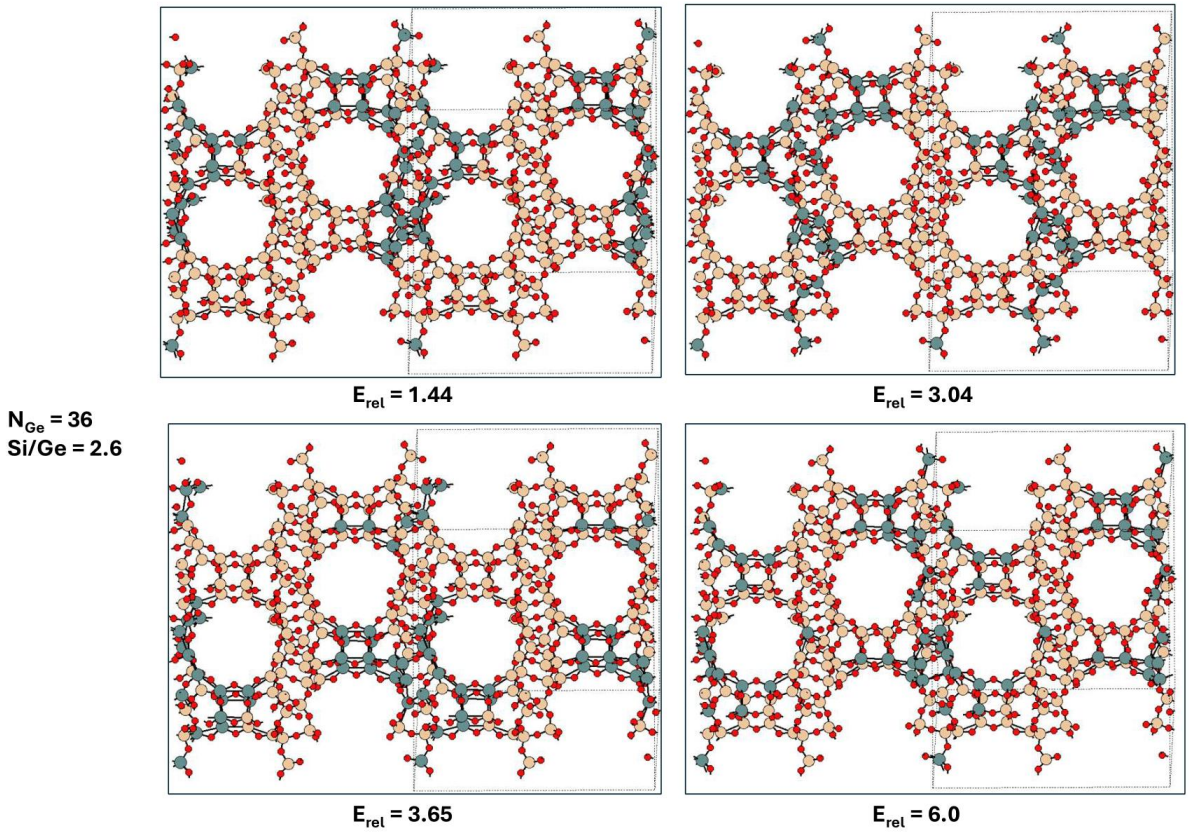


Fig. S47 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for CTH(-A) zeolite (super-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to maximum germanium considered in our simulation. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the c-axis is horizontal, and the b-axis is vertical.

6.4 IWW Supercell

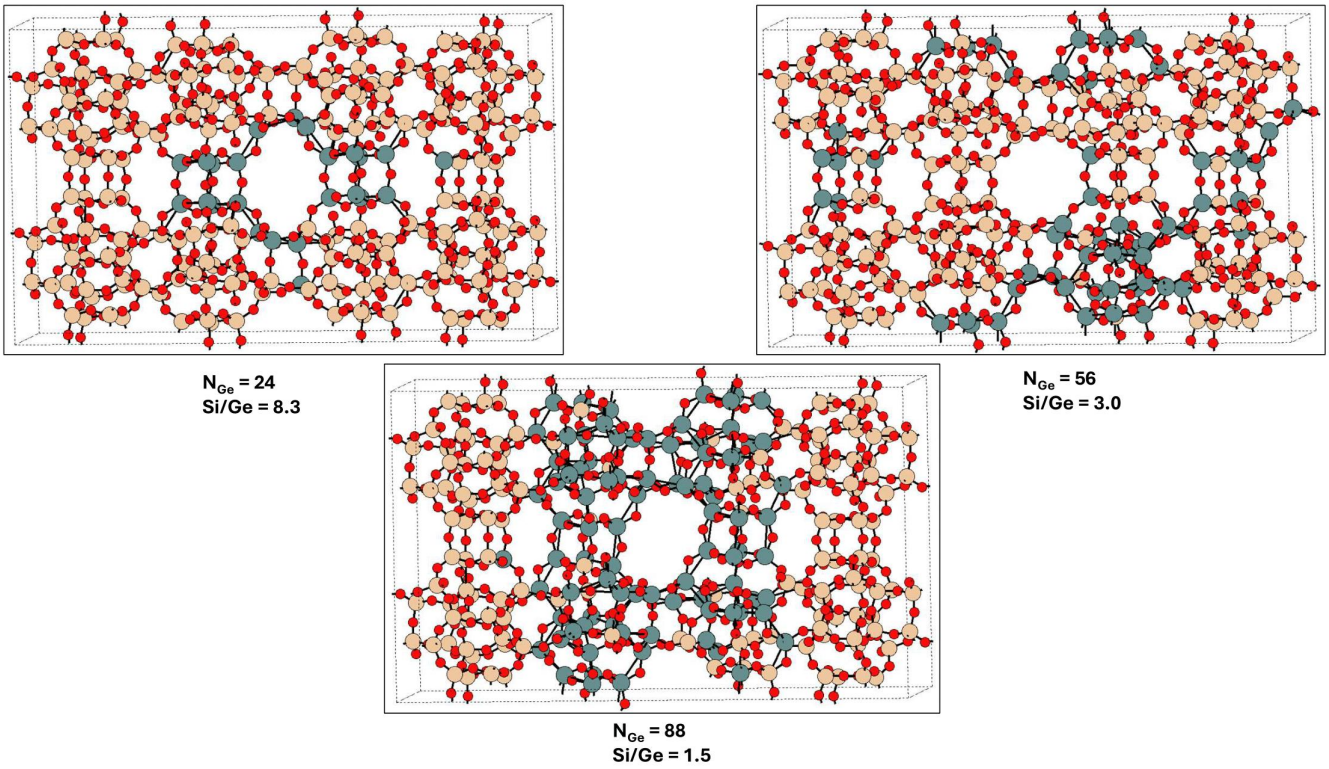


Fig. S48 The GSO structures for IWW zeolite (super-cell model) at various Si/Ge where N_{Ge}, No. of Ge/Unit Cell. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the b-axis is horizontal, and the c-axis is vertical.

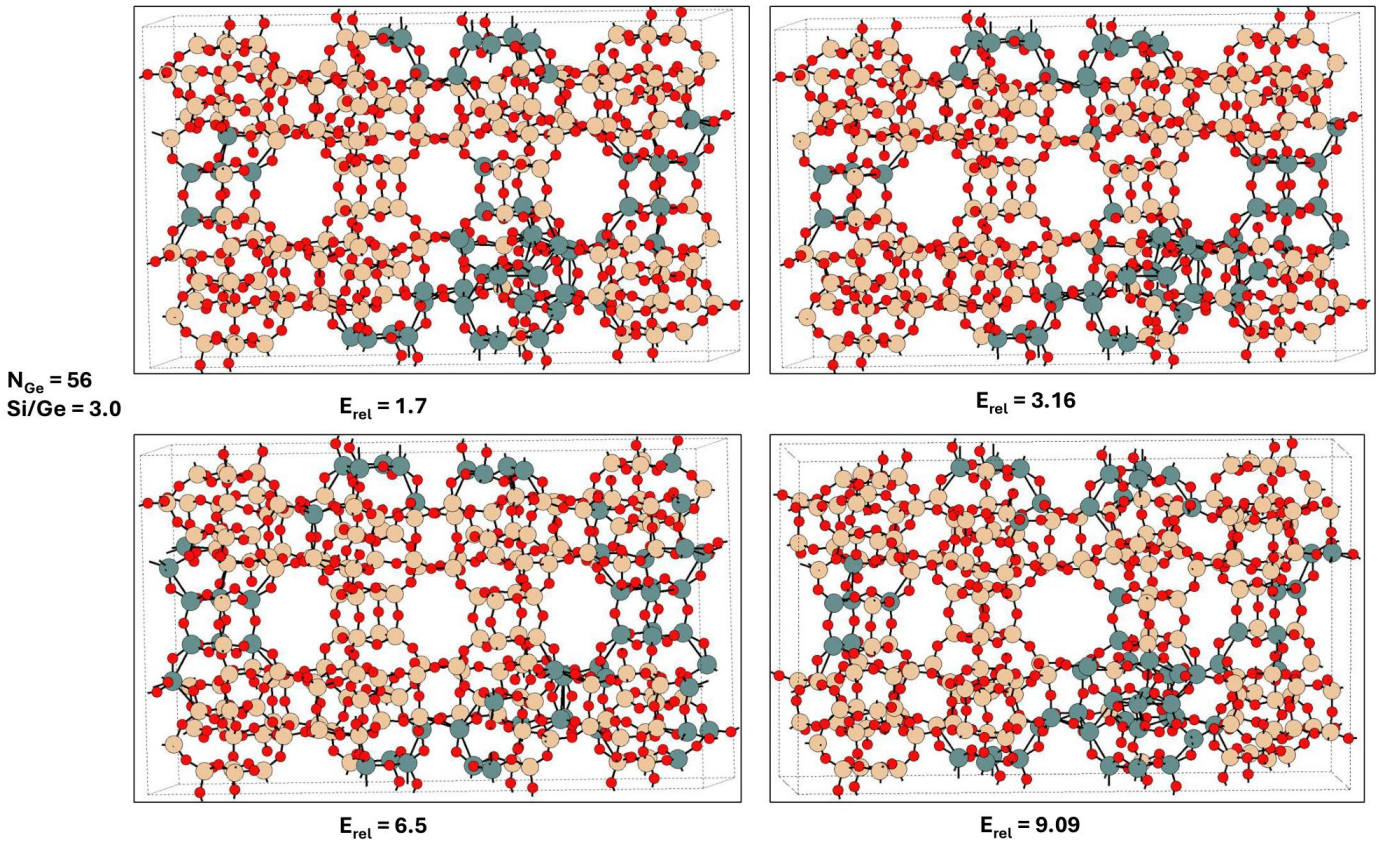


Fig. S49 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for IWW zeolite (super-cell model) where N_{Ge}, No. of Ge/Unit Cell is almost equal to the full-filling of all D4R unit T-sites. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the b-axis is horizontal, and the c-axis is vertical.

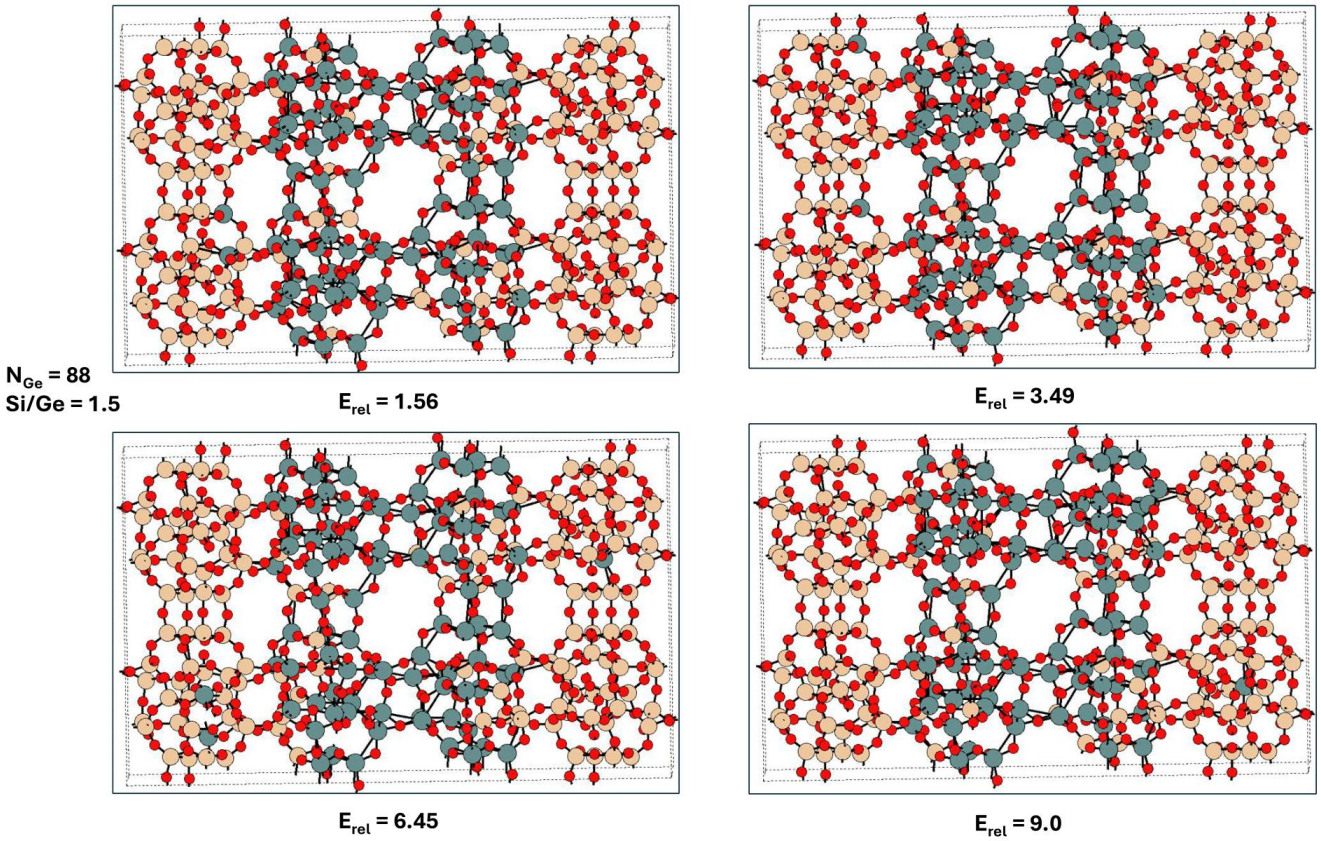


Fig. S50 Representative low-energy structures (with relative energy $E_{\text{rel}} < 10\,\text{kJ/mol}$) for IWW zeolite (super-cell model) where N_{Ge}, No. of Ge/Unit Cell is equal to maximum germanium considered in our simulation. The zeolite structure is viewed along the a-axis, onto the b-c plane. In this orientation, the b-axis is horizontal, and the c-axis is vertical.

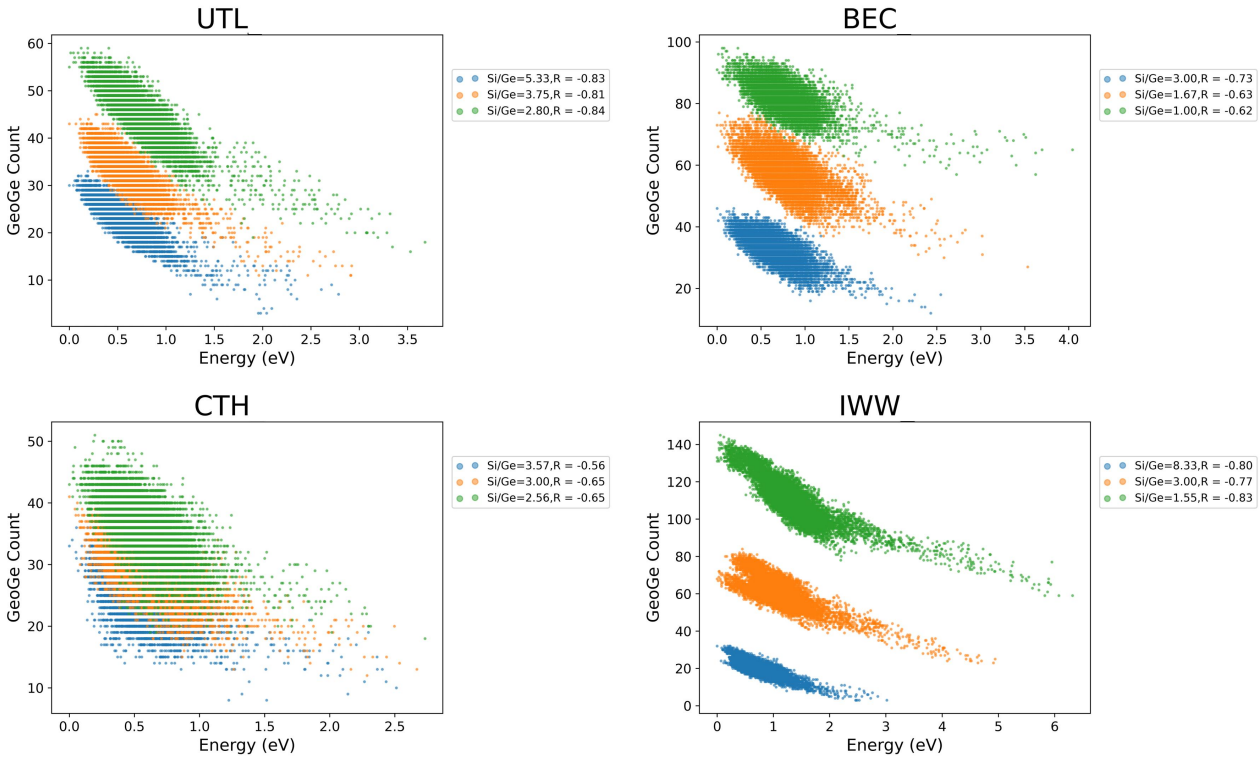


Fig. S51 The correlation between the unique Ge-O-Ge count in a structure and a potential energy of the structure (in eV) for the zeolite topologies (supercells) considered across a range of Si/Ge ratios. The Pearson's correlation coefficients (R)[22] are provided in the legend description.

References

1 J. P. Perdew, K. Burke and M. Ernzerhof, *Physical Review Letters*, 1996, **77**, 3865.

2 Y. Zhang and W. Yang, *Physical Review Letters*, 1998, **80**, 890.

3 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *The Journal of Chemical Physics*, 2010, **132**, 154104.

4 J. M. Wiesenfeld, J. Stone, D. Marcuse, C. Burrus and S. Yang, *Journal of applied physics*, 1987, **61**, 5447–5454.

5 S. Grimme, S. Ehrlich and L. Goerigk, *Journal of Computational Chemistry*, 2011, **32**, 1456–1465.

6 A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus *et al.*, *Journal of Physics: Condensed Matter*, 2017, **29**, 273002.

7 K. Schütt, P. Kessel, M. Gastegger, K. Nicoli, A. Tkatchenko and K.-R. Müller, *Journal of Chemical Theory and Computation*, 2018, **15**, 448–455.

8 R. Fletcher, *Practical methods of optimization*, John Wiley & Sons, 2000.

9 Y.-x. Yuan, *IMA Journal of Numerical Analysis*, 1991, **11**, 325–332.

10 R. T. Rigo, S. R. Balestra, S. Hamad, R. Bueno-Perez, A. R. Ruiz-Salvador, S. Calero and M. A. Camblor, *Journal of Materials Chemistry A*, 2018, **6**, 15110–15122.

11 A. Rojas and M. A. Camblor, *Angewandte Chemie (International Ed. in English)*, 2012, **51**, 3854–3856.

12 A. Corma, M. T. Navarro, F. Rey, J. Rius and S. Valencia, *Angewandte Chemie*, 2001, **113**, 2337–2340.

13 G. Sastre, J. A. Vidal-Moya, T. Blasco, J. Rius, J. L. Jordá, M. T. Navarro, F. Rey and A. Corma, *Angewandte Chemie International Edition*, 2002, **41**, 4722–4726.

14 M. Fischer, *The Journal of Physical Chemistry C*, 2018, **123**, 1852–1865.

15 Y. Wang, J. Song and H. Gies, *Solid state sciences*, 2003, **5**, 1421–1433.

16 G. Sastre and J. D. Gale, *Chemistry of Materials*, 2003, **15**, 1788–1796.

17 H. Li and O. Yaghi, *Journal of the American Chemical Society*, 1998, **120**, 10569–10570.

18 J.-L. Paillaud, B. Harbuzaru, J. Patarin and N. Bats, *Science*, 2004, **304**, 990–992.

19 T. Conradsson, M. Dadachov and X. Zou, *Microporous and mesoporous materials*, 2000, **41**, 183–191.

20 G. Sastre and A. Corma, *The Journal of Physical Chemistry C*, 2010, **114**, 1667–1673.

21 S. O. Odoh, M. W. Deem and L. Gagliardi, *The Journal of Physical Chemistry C*, 2014, **118**, 26939–26946.

22 J. Lee Rodgers and W. A. Nicewander, *The American Statistician*, 1988, **42**, 59–66.

23 International Zeolite Association, *Database of Zeolite Structures*, `http://www.izastructure.org/databases/`, Accessed 2021-03-02.