

## Supplementary Information

for

# Implementing Reactivity in Molecular Dynamics Simulations with Harmonic Force Fields

by

Jordan J. Winetrout,<sup>1,2</sup> Krishan Kanhaiya,<sup>1,2,3</sup> Joshua Kemppainen,<sup>4</sup> Pieter J. in 't Veld,<sup>5</sup> Geeta Sachdeva,<sup>6</sup> Ravindra Pandey,<sup>6</sup> Behzad Damirchi,<sup>7</sup> Adri van Duin,<sup>7</sup> Gregory M. Odegard,<sup>4</sup>  
Hendrik Heinz<sup>1,2\*</sup>

<sup>1</sup> Department of Chemical and Biological Engineering, University of Colorado at Boulder,  
Boulder, CO 80309, USA

<sup>2</sup> Materials Science and Engineering Program, University of Colorado at Boulder, Boulder, CO  
80309, USA

<sup>3</sup> Insitute of Physics, Ruhr University Bochum, Universitätsstrasse 150, 44801 Bochum, Germany

<sup>4</sup> Department of Mechanical Engineering - Engineering Mechanics, Michigan Technological  
University, Houghton, MI 49931, USA

<sup>5</sup> BASF SE, Molecular Modeling & Drug Discovery, 67056 Ludwigshafen, Germany

<sup>6</sup> Department of Physics, Michigan Technological University, Houghton, MI 49931, USA

<sup>7</sup> Department of Mechanical Engineering, Pennsylvania State University, University Park, PA  
16802, USA

\* Corresponding author: [hendrik.heinz@colorado.edu](mailto:hendrik.heinz@colorado.edu)

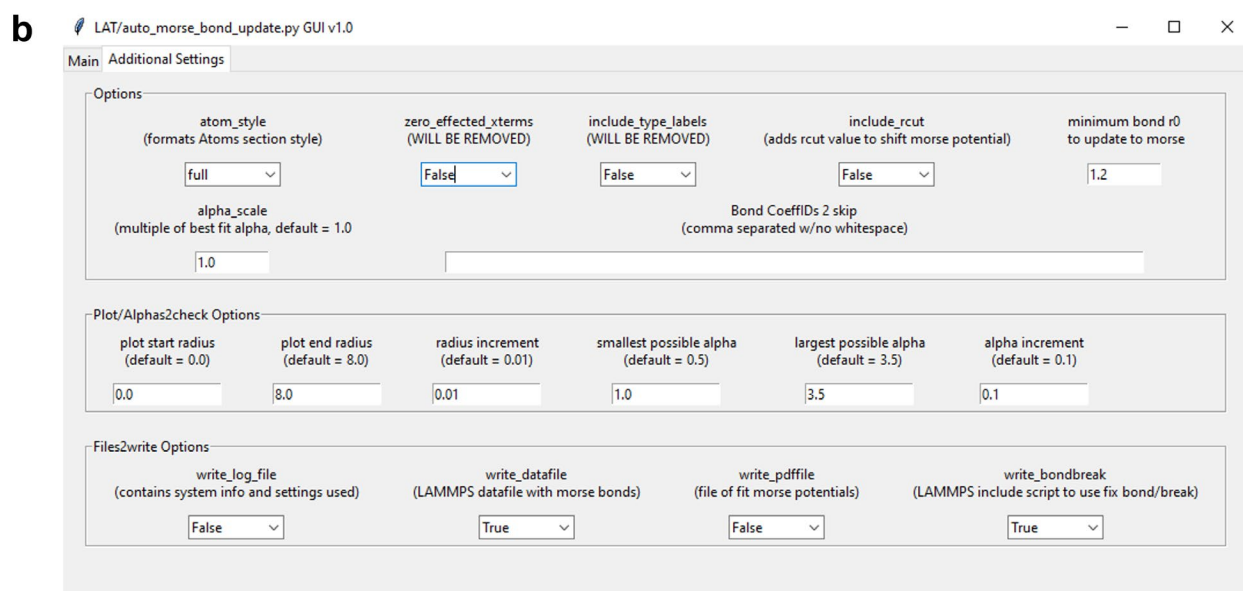
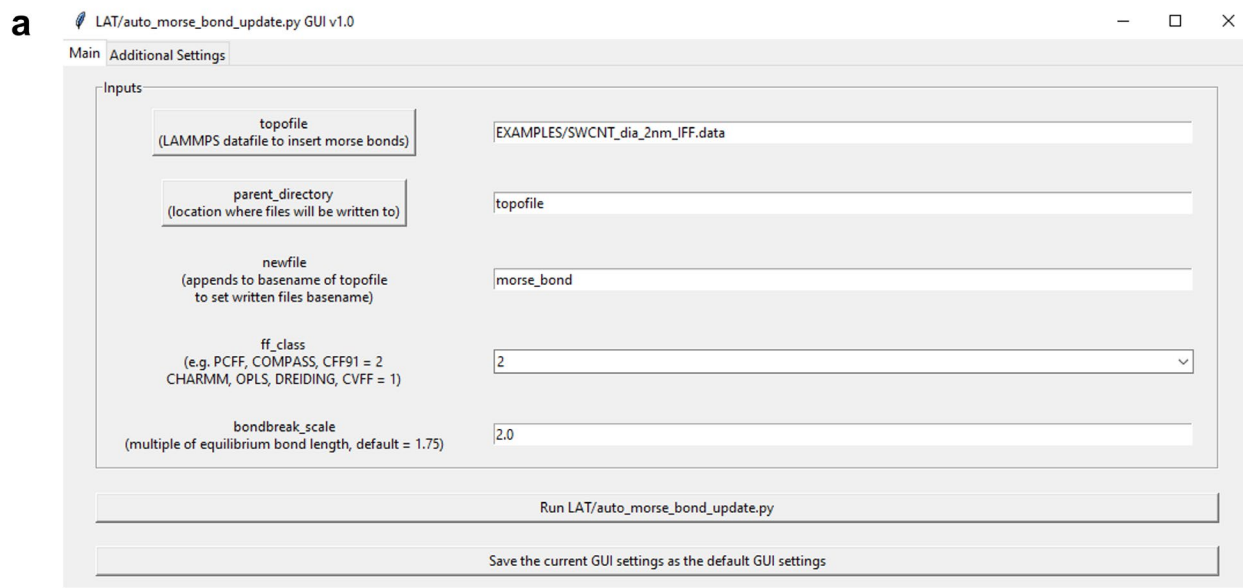
**This PDF file includes:**

Supplementary Figures (pages: S3-S21); Figures S1 to S11)

Supplementary Tables (pages: S22-S24); Tables S1 to S2)

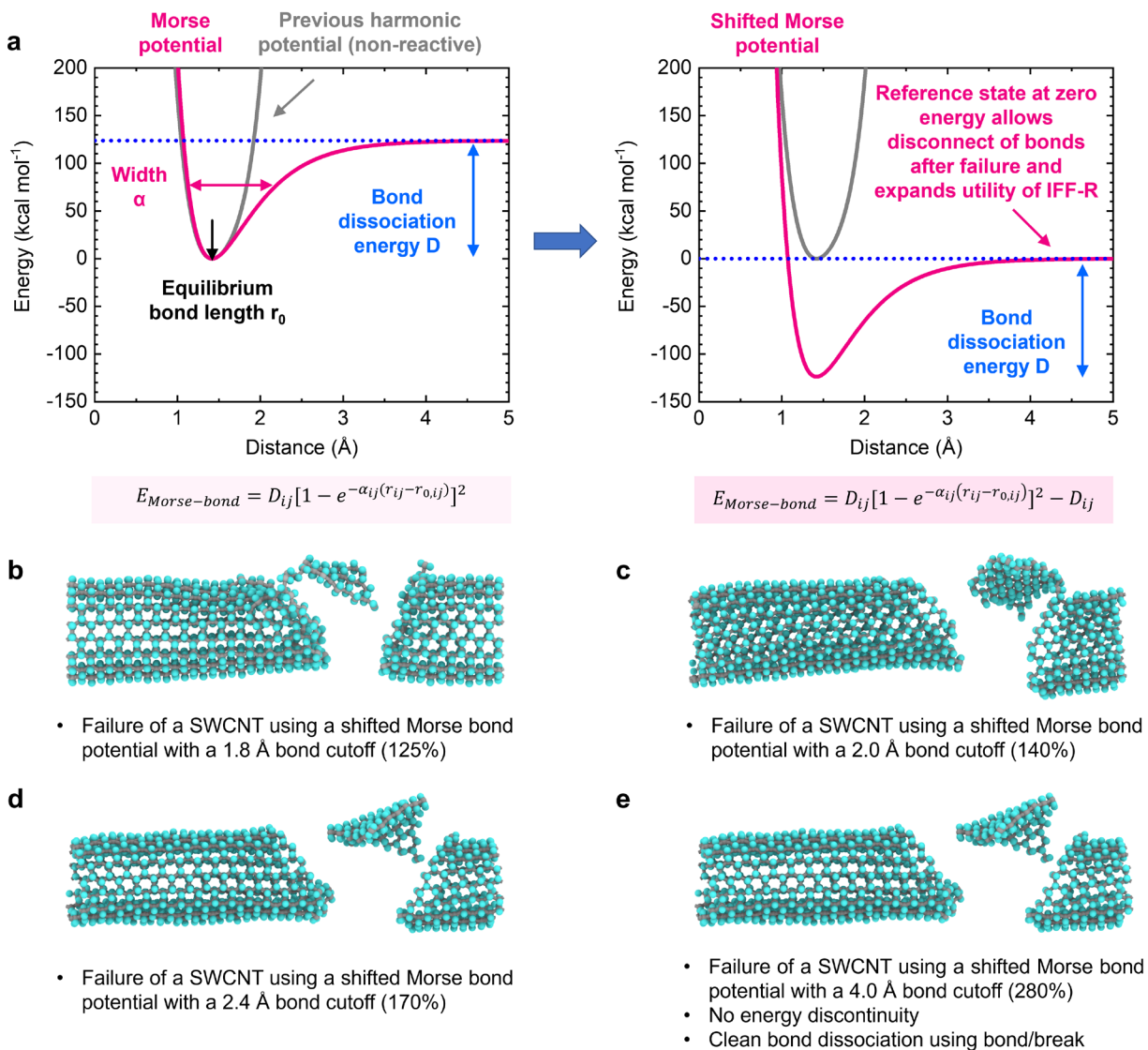
Supplementary Discussion including Detailed Computational Methods (pages S25-S51)

Supplementary References 1-61: (pages: S51-S54)



**Figure S1.** Snapshot of the graphical user interface (GUI) for Morse bond conversion from a classical, harmonic force field (`auto_morse_bond_update.py`). The input to the code is a LAMMPS data file with classical bond parameters. The outputs include a data file with Morse bond parameters, a LAMMPS input script file containing “fix bond/break” commands for the newly assigned Morse bonds, a pdf file showing Morse bond parameterization graphically, and a logfile with system information that can be useful for tracking parameter assignment. (a) The main tab

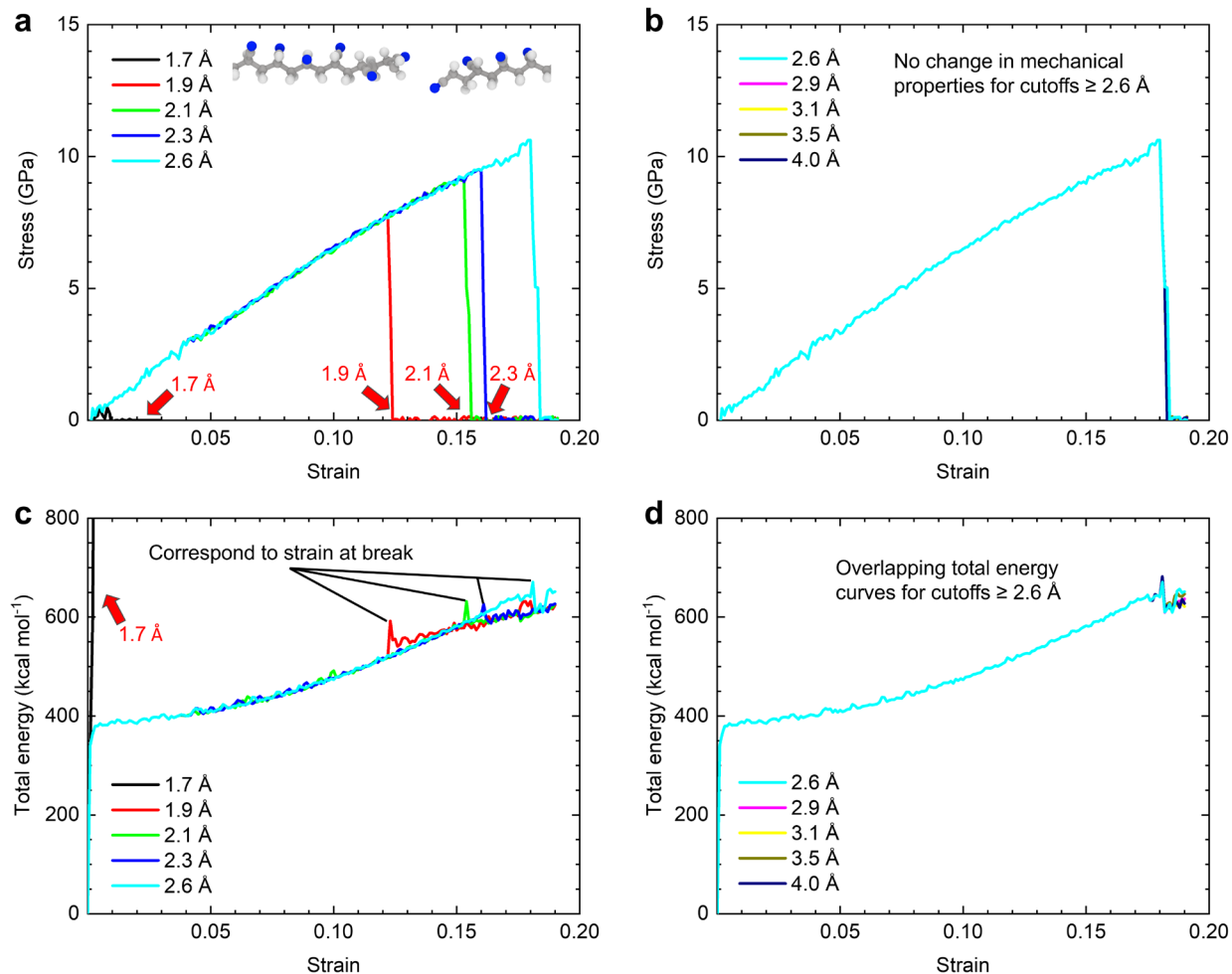
used for automatic Morse bond conversion. The “topofile” button is used for selecting the LAMMPS data file for Morse bond conversion. The “parent\_directory” button is the default path that determines the directory for the outputs, and “topofile” is the path of the LAMMPS data file (topofile). If “topofile” is not used for the “parent\_directory” option, the full path to the desired output directory should be given. The “newfile” section is an additional descriptor appended to the “topofile” name to identify the new data file with Morse bonds and avoids overwriting the “topofile” if it is in the same output directory. The `auto_morse_bond_update.py` code will work with class I and class II force fields, and the “ff\_class” section needs to designate 1 or 2 as acceptable inputs, accordingly. The “bondbreak\_scale” is used cutoff distances for the new Morse bonds (used with the “fix bond/break” command), and the unit is the multiple of the equilibrium bond length. We recommend this parameter to be 2.0, e.g., creating a bond/break cutoff for a C-C single bond with  $r_0 = 1.54 \text{ \AA}$  at  $3.08 \text{ \AA}$ . (b) Image of the “Additional Settings” tab. This tab is included so the user has customizability over the outputs from the `auto_morse_bond_update.py` code. The “Options” section provides capabilities for the user to modify the information written to the output datafile. The “Plot/Alphas2check Options” section allows the user to change variables associated with the automatic parameterization of a Morse bond. Finally, the “Files2write Options” section allows the users to specify which files should appear in the output directory with “False” meaning the file will not be written to the output directory and “True” meaning the file will be written to the output directory. Using this menu of additional settings is optional.



**Figure S2.** Illustration of the shifted Morse bond potential in IFF-R and implications on failure (see Supplementary Software “code for custom LAMMPS option user morse-2Aug23”). The Morse potential benefits from energy conservation, i.e., a shift to zero energy upon bond dissociation. (a) The default Morse bond potential (left) and the shifted Morse bond potential (right) for graphitic carbon-carbon bonds with an equilibrium bond length of 1.42 Å. The unmodified Morse bond potential has an energy minimum at  $r_{0,ij}$  and 0 kcal mol<sup>-1</sup> and a high potential energy upon bond dissociation, which causes an energy discontinuity (energy drop) when

elongated Morse bonds are disconnected or reassigned, respectively. In comparison, the shifted Morse bond potential has an energy minimum at  $r_{0, ij}$  and  $-D_{ij}$ . Using a shift so that the bond energy approaches 0 kcal/mol as the bond dissociates avoids energy discontinuities in molecular dynamics simulations arising from the Morse potential during bond scission. At the same time, the shift does not change the force as a function of distance and has no visible effect on the failure mechanism.

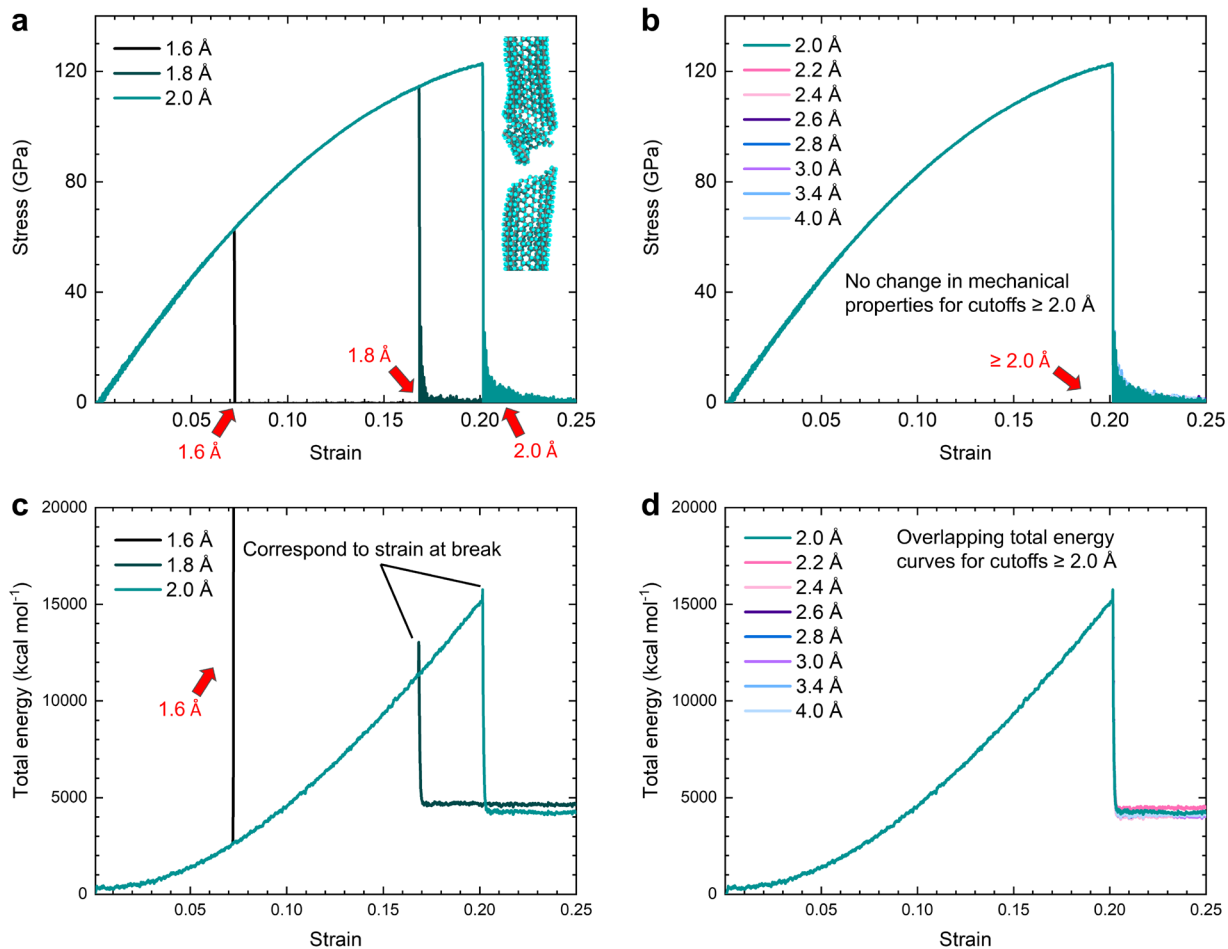
(b-e) Another key parameter is the bond cutoff distance. If the bonds are not disconnected at a certain distance past dissociation, “ghost” contributions from remaining angle, dihedral, and other energy contributions taint the potential energy. On the other hand, cutoff distances too close to the equilibrium bond length are unphysical. Examples of the fracture pattern of a SWCNT are shown using the shifted Morse potential with bond cutoff distances of 1.8 Å (b), 2.0 Å (c), 2.4 Å (d), and 4.0 Å (e). The simulation protocol was the same. The failure mechanism was affected when the cutoff distances were chosen too small and became independent of the cutoff distance when chosen larger than 2.4 Å, equal to  $\sim 170\%$  of the equilibrium bond length.



**Figure S3.** Stress-strain curves of a polyacrylonitrile chain and the total energy for 9 different bond cutoffs ( $R_{max}$ ) using a Morse potential shifted by  $-D_{ij}$ . The bond/break LAMMPS command was used during a simulation of the tension response. (a) Stress-strain curve for  $R_{max}$  values from of 1.7 Å to 2.6 Å, equal to 111% to 170% of the original C-C bond length of 1.53 Å. The inset in the upper left shows the polyacrylonitrile chain after failure using a cutoff distance  $R_{max} \geq 2.6$  Å. Cutoffs below 2.6 Å cause premature failure, particularly for a short cutoff at 1.7 Å (arrows in red). (b) Stress-strain curve for  $R_{max}$  values from 2.6 Å to 4.0 Å, equal to 170% to 261% of the original C-C bond length of 1.53 Å. All bond cutoffs in this range show almost identical failure and correspond to equilibrium conditions in the IFF-R simulation. (c) Total energy vs strain curves

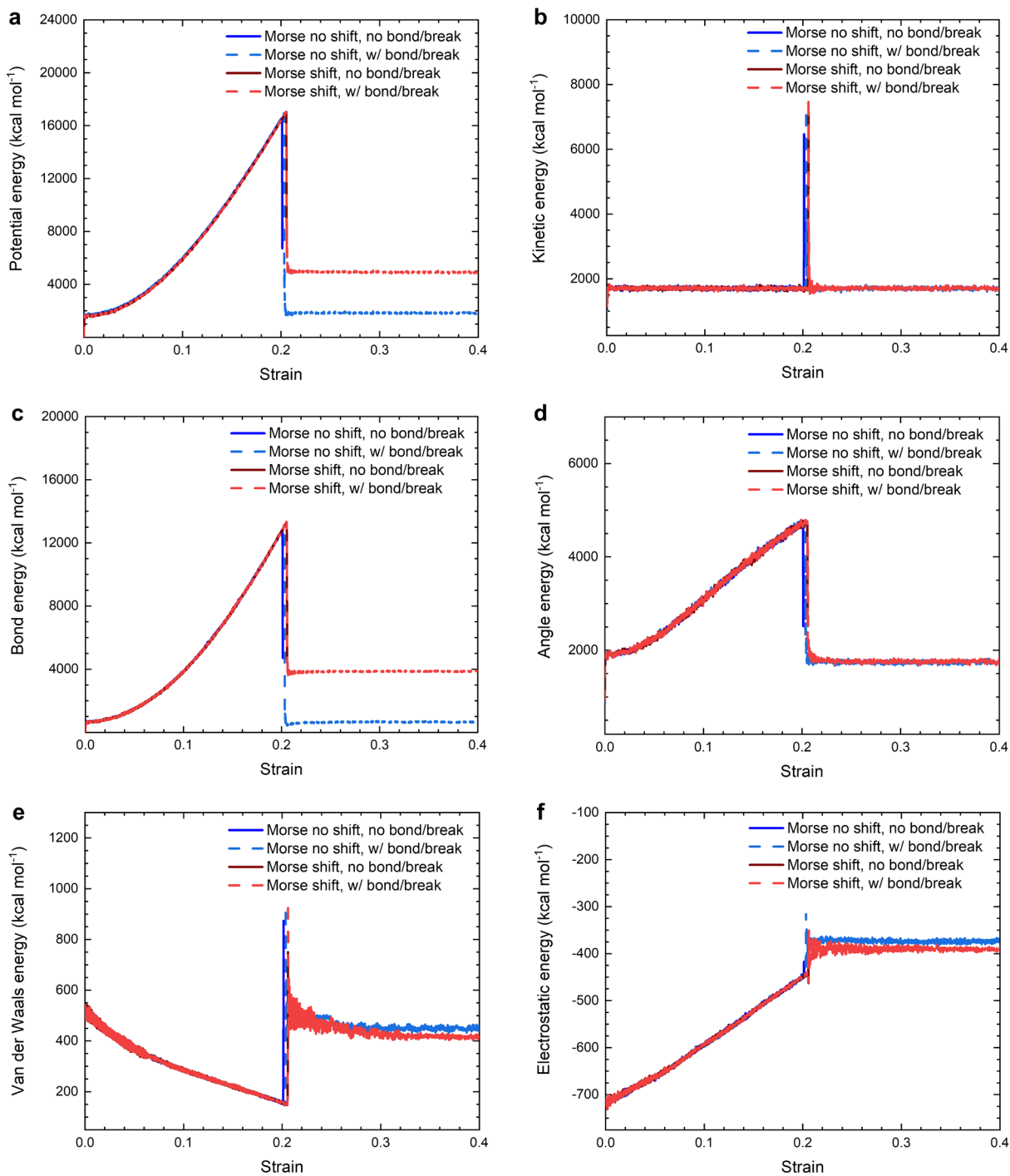
for bond cutoffs from 1.7 Å to 2.6 Å shown in (a). The total energy spikes instantaneously for a  $R_{max}$  of 1.7 Å and shows discontinuities for  $R_{max}$  values up to 2.6 Å, near the respective strain at break. (d) Total energy vs strain curves for bond cutoffs from 2.6 Å to 4.0 Å shown in (b). The total energy curves are identical, indicating equilibrium bond dissociation. Small, similar discontinuities at break are seen. Overall, a minimum bond cutoff of at least 170% of the equilibrium bond distance of C-C single bonds is recommended.





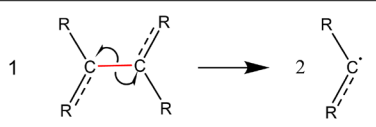
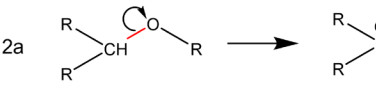
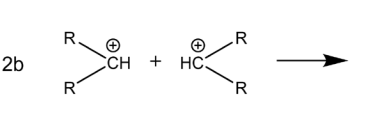
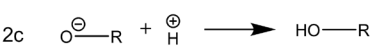
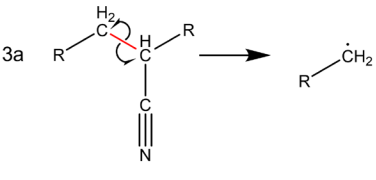
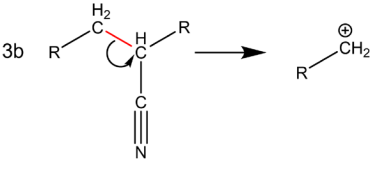
**Figure S4.** Stress-strain curves and the total energy of a single wall carbon nanotube (SWCNT) for 10 different bond cutoffs ( $R_{max}$ ) using a shifted Morse potential (by  $-D_{ij}$ ). The bond/break LAMMPS command was used during a tensile simulation. (a) Stress-strain curve for  $R_{max}$  values from of 1.6 Å to 2.0 Å, equal to 113% to 141% of the original  $C_{ar}-C_{ar}$  bond length of 1.42 Å. Cutoffs below 2.0 Å cause premature failure (arrows in red). The inset in the upper left shows a SWCNT after failure with cutoff  $R_{max} \geq 2.0$  Å. (b) Stress-strain curve for  $R_{max}$  values from 2.0 Å to 4.0 Å, equal to 141% to 282% of the original C-C bond length of 1.42 Å. All bond cutoffs in this range show almost identical failure and correspond to equilibrium conditions in the IFF-R simulation. (c) Total energy vs strain curves for bond cutoffs from 1.6 Å to 2.0 Å shown in (a). The total energy shows discontinuities near the respective strain at break for the different  $R_{max}$

values. (d) Total energy vs strain curves for bond cutoffs from 2.0 Å to 4.0 Å shown in (b). The total energy curves are nearly identical, indicating equilibrium bond dissociation. Small, similar discontinuities at break are seen. Overall, a minimum bond cutoff of at least 140% of the equilibrium bond distance of C<sub>ar</sub>-C<sub>ar</sub> aromatic bonds (1.53 Å) is recommended.



**Figure S5.** Energy contributions as a function of strain during simulations of tensile stress-strain curves of a SWCNT using IFF-R Morse bond parameters. The potential energy and bond energy for the default Morse potential and the shifted Morse potential have been normalized to 0 kcal/mol

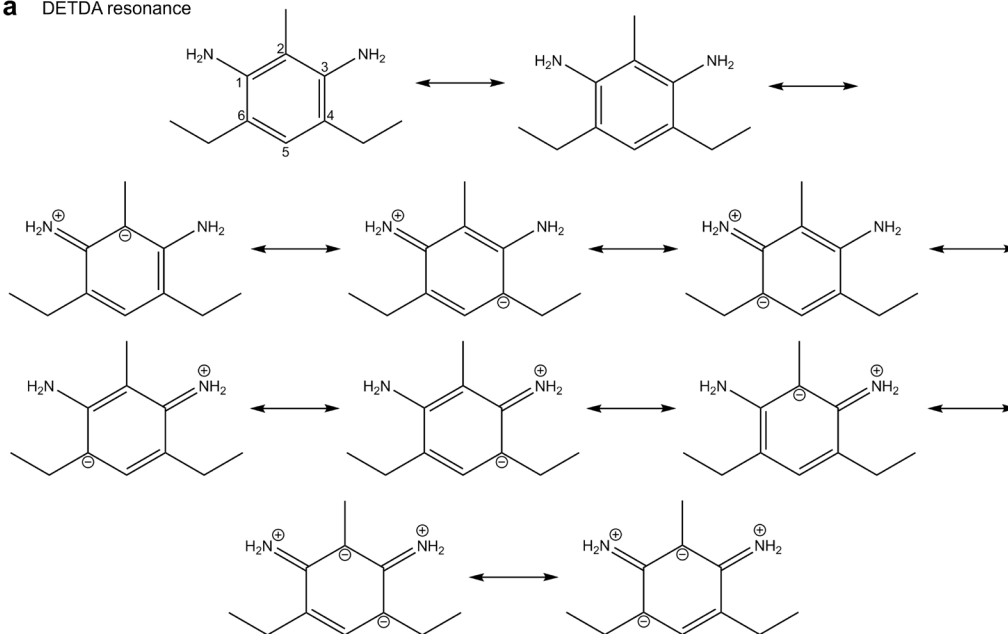
for comparison. Differences between the default Morse potential, the shifted Morse potential (by  $D_{ij}$ ), and the default and shifted Morse potential with bond break ( $R_{max\ Car-Car} = 2.84\ \text{\AA}$ ) can be seen. Contributions from the torsion potential (that stabilizes a flat  $sp^2$  geometry) are not shown. One of the future challenges is the reduction, or elimination, of discontinuities in energy upon bond breaking without adding artificial new parameter space. (a) The total potential energy, (b) the kinetic energy, (c) the bond energy, (d) the angle energy, (e) the van der Waals energy, and (f) the electrostatic energy. The bond energy (a, c) continues to show discontinuities upon dissociation since also neighbor bonds are strained and released in the process of disconnecting the strained Morse bond, as well as angles, torsions, and nonbond interactions in the vicinity of the reaction site (d, e, f). The kinetic energy temporarily increases due to gained mobility (b). The Coulomb energy increases due to repulsion between like charges after bond C-C scission (f). The effects depend on whether new atom types and force field parameters are assigned after bond scission.

Bond type	Dissociation mechanisms	Parameterization recommendations
Graphitic		Update nonbonded parameters for dissociated carbon atoms
Cellulose		Update charge distribution for carbon and oxygen atoms
		Create new chemical bonds with two C atoms redefined as sp <sup>2</sup> /alkene
		Create new chemical bond and redefine R-O <sup>-</sup> oxygen to hydroxyl oxygen atoms
PAN		Update nonbonded parameters for dissociated carbon atoms
		Update charge distributions for dissociated carbon atoms

**Figure S6.** Bond dissociation mechanisms for C<sub>ar</sub>-C<sub>ar</sub> graphitic bonds, C-O bonds in cellulose, C-C bonds in polyacrylonitrile and suggestions to parameterize the newly dissociated atoms (ref. <sup>1</sup>). Proposed cleaved bonds are shown in red and R designates substituent groups not participating in the reactions/dissociation. Various other follow-on reactions of radicals and ions are possible, e.g., with oxygen, water from moisture, and atmospheric gases depending on the reaction conditions. Mechanism 1 indicates homolytic cleavage of C<sub>ar</sub>-C<sub>ar</sub> bonds in graphitic structures and carbon nanotubes into radicals. Mechanism 2a shows the heterolytic cleavage of C-O ether linkages in cellulose. If bond formation can be considered in the simulation, mechanism 2b proposes the stabilization of carbocation intermediates through the formation of a double bond, and mechanism 2c the protonation of alcoholate groups. Mechanism 3a proposes homolytic dissociation of C-C bonds in PAN, and mechanism 3b heterolytic dissociation, both of which could lead to multiple

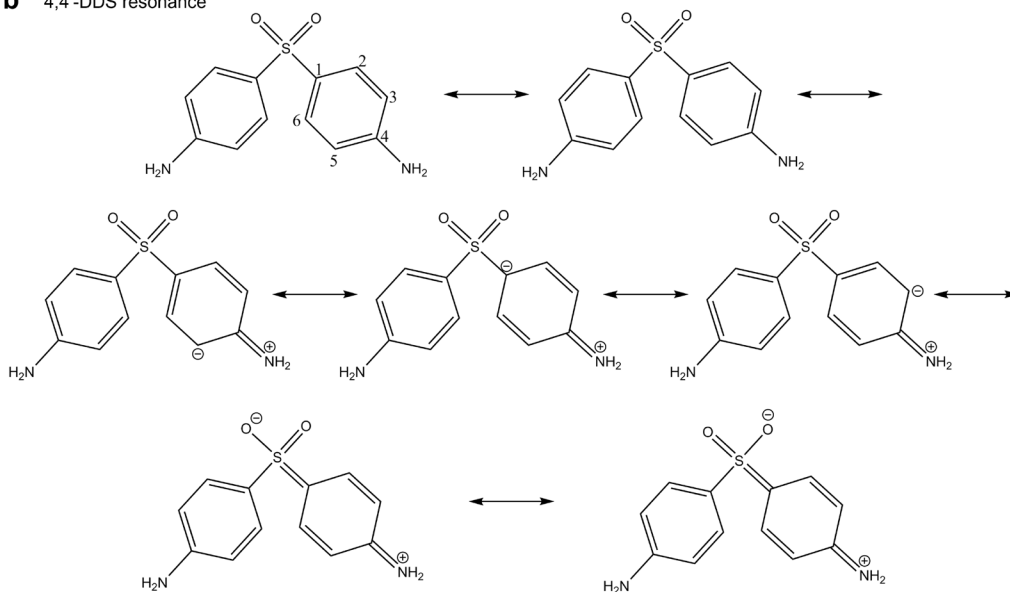
products. All proposed follow-on reactions and parameterization recommendations can be handled via the LAMMPS REACTER package (ref. <sup>2</sup>).

**a** DETDA resonance



Bond	No. of single bonds	No. of double bonds	Double bond character (%)
1, 2	7	3	30
2, 3	7	3	30
3, 4	8	2	20
4, 5	5	5	50
5, 6	5	5	50
6, 1	8	2	20

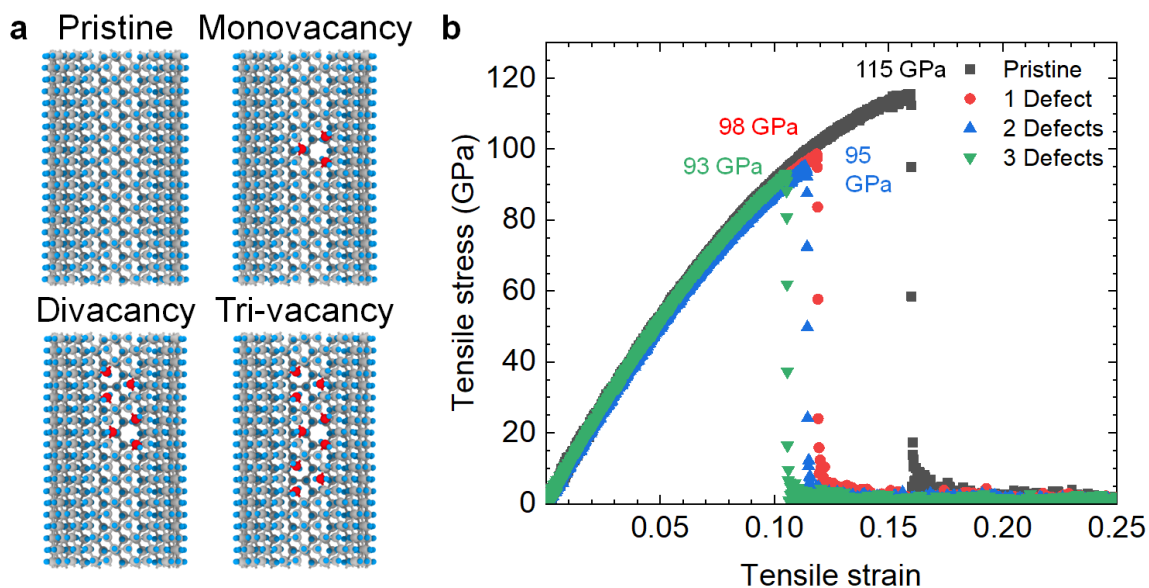
**b** 4,4'-DDS resonance



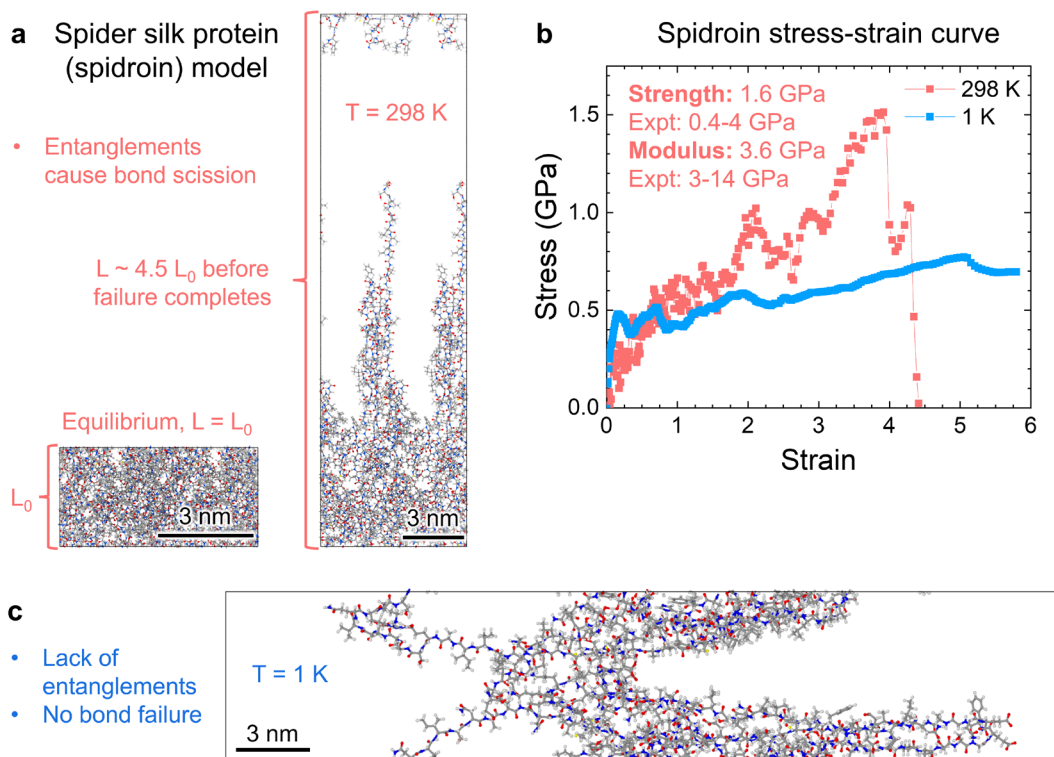
Bond	No. of single bonds	No. of double bonds	Double bond character (%)
1, 2	5	2	28.6
2, 3	2	5	71.4
3, 4	6	1	14.3
4, 5	6	1	14.3
5, 6	2	5	71.4
6, 1	5	2	28.6

**Figure S7.** Resonance structures of (a) DETDA and (b) 4,4'-DDS with atom labels (top) and estimated double bond character of individual bonds. Higher double bond character correlates with increased bond dissociation energy. The percentage of double bond character for each bond was calculated as the fraction of resonance structures showing double bonds. In DETDA, double bond character increases in the order  $3,4 < 1,2 < 5,6$ . In 4,4'-DDS, the double character increases in the order  $3,4 < 1,2 < 2,3$  with high double bond character of the 2,3 and 5,6 bonds (>70%) due to a push-pull effect from the p-situated amino and sulfoxide groups.



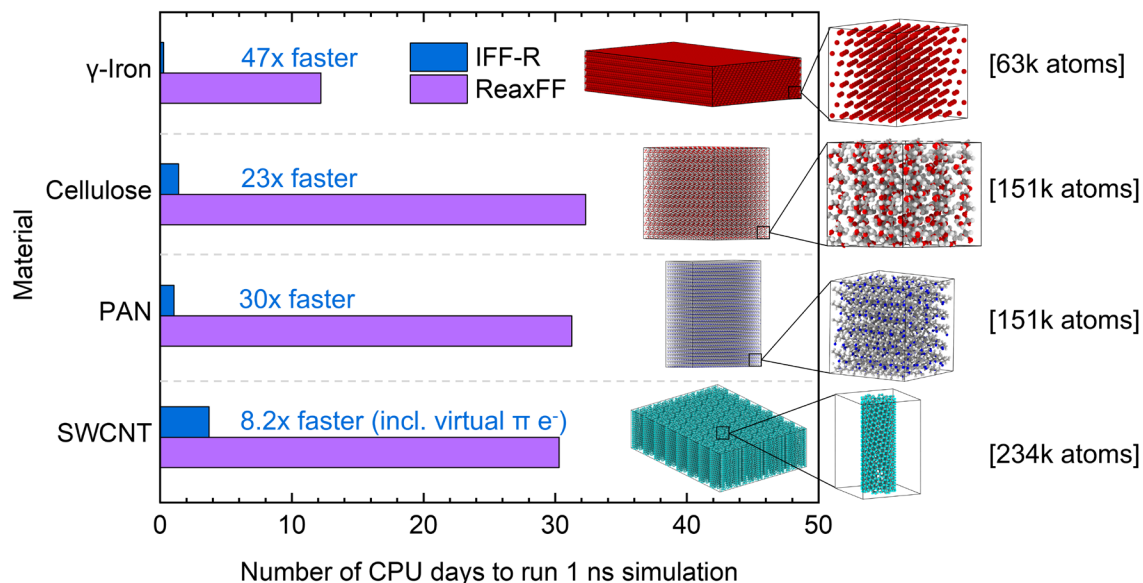


**Figure S8.** Demonstration of the variability of the stress-strain curves and failure mechanisms of single-wall carbon nanotubes containing defects as seen in MD simulations with IFF-R. (a) Models of 2.0 nm diameter (15,15) SWCNTs with increasing number of missing atoms from 0 (pristine) to 3 (tri-vacancy). Sites of missing atoms are highlighted in red. (b) Stress-strain curves for the carbon nanotubes in (a). A limit for the influence of missing atoms on the reduction of tensile strength and ultimate strain can be seen. A single missing atom from a pristine SWCNT structure causes the greatest reduction in tensile strength of -14.7%, i.e., from 115 GPa to 98 GPa. Subsequent removal of atoms leads to reductions of -3.1% and -2.1%, respectively.

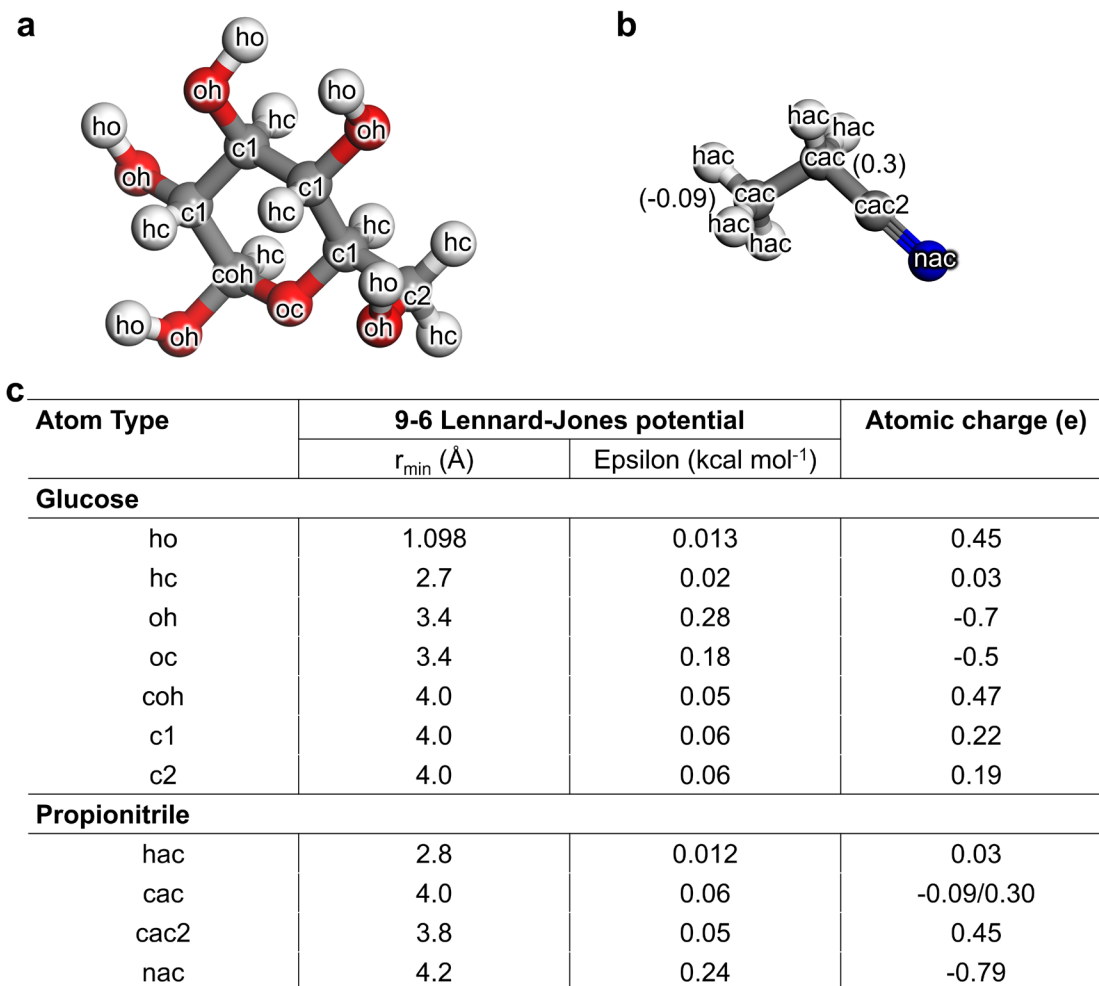


**Figure S9.** Use of CHARMM36m parameters<sup>3</sup> with IFF-R Morse bonds for the simulation of tensile failure of the spider silk protein spidroin. (a) A model of spidroin consisting of two protein strands in equilibrium (left) and at an elongation of more than 4 times the original box length (right) after failure at a temperature of 298 K. The dynamics at 298 K led to the formation of entanglements of the chains, which affect the failure mechanism. (b) Stress-strain curve of the spidroin model obtained by reactive molecular dynamics simulations. In the chosen model configuration, spidroin showed a tensile strength of 1.6 GPa and a Young's modulus of 3.6 GPa. The values are in the range of the experimentally determined tensile strength of 0.4 to 4.0 GPa and the experimentally measured tensile modulus of 3 to 14 GPa for spider silk (refs. <sup>4-7</sup>). No modifications of CHARMM36m parameters<sup>8</sup> for spidroin, other than the introduction of IFF-R Morse bonds were made to obtain the data reported. (c) The response of spidroin to tensile stress at a lower temperature of 1 K. Entanglements of the protein strands were not observed due to the

limited mobility of the amino acids at 1 K and led to different mechanical properties. Using the same starting configuration as at 298 K, the two spidroin molecules separated during the tensile simulation, the stress response was dominated by nonbonded interactions, and bond breaking did not occur. At either temperature, the mechanical properties are also influenced by the nanoscale morphology, relaxation time, and strain rate, which were not further examined here.



**Figure S10.** Comparison of CPU time needed for the simulation of stress-strain curves using IFF-R and ReaxFF for large model sizes of 63,000 to 234,000 atoms. CPU time and snapshots at failure are shown for a crystal of  $\gamma$ -iron, a crystal of cellulose I $\beta$  (ref. <sup>9</sup>), a crystal of syndiotactic poly(acrylonitrile) (PAN), and bundle of single-walled carbon nanotubes (SWCNT) of 1 nm diameter using IFF-R (PCFF-R in case of cellulose). The materials are the same as in Figure 4, however, scaled 50-100x larger for  $\gamma$ -iron and the polymers, and  $\sim$ 20x larger for the CNT bundle. Full scale images of the model systems are shown in the chart area and close-up images on the right-hand side of the graph. The benchmark simulations were carried out on 24 CPUs, using 1000 steps of molecular dynamics simulations with a constant strain rate in the tensile direction, and extrapolated for 1 ns duration. The NPT ensemble was employed with pressure control in  $x$  and  $y$  directions, and no pressure control in the tensile ( $z$ ) direction. IFF-R completes 1 ns molecular dynamics simulation using 24 CPUs in 6 hours to 4 days for all 4 materials. The ReaxFF simulations would finish between 12 days and 1 month. The efficiency of IFF-R is about the same as for small model systems, including a slight decrease for cellulose I $\beta$  and  $\gamma$ -iron crystals, and a slight increase for PAN and CNTs.



**Figure S11.** Models, force field types, and nonbonded parameters for  $\alpha$ -D-glucose and propionitrile. (a, b) The structure of  $\alpha$ -D-glucose and propionitrile with atom types (force field types). The model of (b) propionitrile also lists the charges on the carbon backbone. (c) Table of force field types, 9-6 Lennard-Jones parameters, and atomic charges consistent with dipole/multipole moments. The parameters for  $\alpha$ -D-glucose are PCFF-R and include updated atomic charges. The parameters for propionitrile are IFF-R (higher quality).

**Table S1.** Morse parameters for selected bonds in IFF-R and the expected range for any chemistry (bottom two lines). Origin of reference data: Experimental data for equilibrium bond lengths are available in ref. <sup>10</sup> and for bond dissociation energies in refs. <sup>10-13</sup>. Data for all types of carbon-carbon bonds additionally consider ref. <sup>14</sup>, and the data for carbon-carbon bonds in graphite and carbon nanotubes also consider ref. <sup>15</sup>. The range of bond energies across a broad range of chemistries from refs. <sup>10-14</sup> can be significantly narrowed down once the specific chemical environment of the bonds is considered (<3% uncertainty).

Bond type	Equilibrium bond length $r_{0,ij}$ (Å)		Width $\alpha$ (Å <sup>-1</sup> ) <sup>b</sup>	Bond dissociation energy $D_{ij}$ (kcal/mol)	
	IFF-R	Experiment <sup>1</sup> <sub>0</sub>		IFF-R	Experiment <sup>10-13</sup>
C <sub>ar</sub> – S (sulfone)	1.80	1.77	1.9	75	75-85
C <sub>ar</sub> – N (amine)	1.39	1.38	2.0	110	100-110
C – C (single bond)	1.53	1.54	2.2	85	70-90
C <sub>ar</sub> – C <sub>ar</sub> (benzene)	1.39	1.38	2.2	150	~155
C <sub>ar</sub> – C <sub>ar</sub> (graphitic and CNTs)	1.42	1.42	2.4	124	108-140
C – O (ether) <sup>a</sup>	1.42	1.43	2.0	85	81-98
C = C (double bond) <sup>a</sup>	1.34	1.34	1.9	152	140-172
Typical range	0.7-2.8	0.7-2.8	1.7-2.4	0-250	0-250
Typical uncertainty	<1%	<1%	<0.1	<10%	<10%

<sup>a</sup> The Morse bond parameters for C=C double bonds and C-O ether bonds are generic and require fine-tuning according to specific chemical environments (see examples in Figure 2). <sup>b</sup> The alpha value may vary due to substituent effects not specifically included here.

**Table S2.** Comparison of surface and vaporization energies for graphite, propionitrile, and  $\alpha$ -D-glucose using experimental data, computational data from IFF-R/PCFF-R, and two different ReaxFF parameter sets by Damirchi et al. (ref. <sup>16</sup>) and Liu et al. (ref. <sup>17</sup>).

<b>Material</b>	<b>Expt</b>	<b>IFF-R</b>	<b>Dev (%)</b>	<b>ReaxFF (Damirchi)</b>	<b>Dev (%)</b>	<b>ReaxFF (Liu)</b>	<b>Dev (%)</b>
<b>Surface energy (mJ/m<sup>2</sup>)</b>							
Graphite	186 $\pm$ 2.0 <sup>18,19</sup>	188 $\pm$ 1.9	+1.1	207 $\pm$ 1.9	+11	242 $\pm$ 2.0	+30
<b>Vaporization energy of monomer (kJ/mol)</b>							
Propionitrile	36.1 $\pm$ 1.6 <sup>13</sup>	37.0 $\pm$ 0.35	+2.5	29 $\pm$ 0.8	-20	47 $\pm$ 0.9	+30
$\alpha$ -D-glucose	181.7 $\pm$ 5.0 <sup>20</sup>	176 $\pm$ 0.33	-3.1	118 $\pm$ 2.5	-35	169 $\pm$ 1.6	-7.0



## S1. Detailed Computational Methods

**S1.1. Scans of Bond Dissociation Energies.** *Model Building.* The amine molecules were aligned in the  $x$ - $y$  plane of the simulation cell and the bonds of interest aligned along the  $x$ -axis to easily modify the interatomic distance. The bond length between atoms  $i$  and  $j$  was elongated with a step-size of 0.1 Å or less while constraining the positions of all other atoms.

*DFT (Density Functional Theory) and MP2 Protocol.* DFT calculations use the B3LYP gradient-corrected hybrid functional,<sup>21,22</sup> and MP2 calculations were based on Hartree-Fock theory with added electron correlations up to second order perturbation theory. MP2 differs from DFT in including corrections to the wavefunction, providing more accurate bond scans than DFT to derive IFF-R parameters. The Gaussian 2016 program<sup>23</sup> was used and employed an all-electron Gaussian basis set to describe electronic configurations of molecules. For the constituent atoms of amine monomers, the standard split-valence polarized double-zeta basis sets 6-31G (d, p) were used. At each point of the energy surface, the convergence criteria for the RMS density matrix and total energy were  $10^{-8}$  and  $10^{-6}$  eV, respectively. No periodic boundary conditions were applied to calculate bond dissociation energies of the molecules. For simplicity, the energy of the equilibrium molecular configuration was set to zero.

*Molecular Mechanics Protocol.* Simulations of bond dissociation curves using IFF-R, as well as the ReaxFF parameter set by Singh et al.<sup>24</sup> were carried out by a series of energy minimizations in LAMMPS using the Polak-Ribiere version of the conjugate gradient algorithm.<sup>25</sup> The minimization stopping criteria were set to a  $10^{-4}$  energy tolerance, or  $10^{-6}$  kcal/(mol·Å) force tolerance, within a maximum of 10,000 steps. During energy minimization at each bond distance, the positions of all atoms were constrained except for the two atoms  $i$  and  $j$  involved in bond dissociation. The equilibrium energies of the molecule were recorded as a function of bond length

and normalized to zero at the distance of lowest energy.

### **S1.2. Simulation of Tensile Modulus and Strength. *Model Building and Parameter Choices.***

A model of an (8, 8) armchair single-wall carbon nanotube with a diameter of 1.085 nm and a length of 4.92 nm (i.e., 20 repeat units in length) was created in Materials Studio. Virtual  $\pi$ -electrons were added for stress-strain simulations with IFF and IFF-R,<sup>26</sup> and removed for simulations with ReaxFF (which does not explicitly represent such features).

A model of crystalline, syndiotactic PAN was created using the “Build Homopolymer” tool in Material Studio. The model matched the experimentally known density<sup>27</sup> of 1.16 g/cm<sup>3</sup> and assumed a rectangular crystalline supercell with parallel chains and lattice dimensions of 3.12 nm x 3.07 nm x 3.42 nm. No experimental information on unit cell parameters and atomic positions, e.g., from X-ray diffraction, was available for syndiotactic 100% crystalline PAN.

A model of crystalline cellulose I $\beta$  was created using the structure built by Nishiyama et al. based on X-ray and neutron diffraction data.<sup>9</sup>

A unit cell for  $\gamma$ -iron was built using the IFF model database with initial estimated density of 8.18 g/cm<sup>3</sup> and expanded into a supercell of the dimensions 1.77 nm x 2.18 nm x 1.85 nm.<sup>28</sup>  $\gamma$ -iron is a suitable model for stainless steel, typically including minor components of other alloying elements (Cr, Ni, C, Mo, Mg). Otherwise, pure iron assumes a bcc structure of  $\alpha$ -iron at room temperature.<sup>29</sup> The force field for simulations was chosen as IFF and IFF-R with 9-6 Lennard-Jones options.<sup>30,31</sup>

We further carried out simulations with ReaxFF for comparison using the same models. Stress-strain simulations of the SWCNT, PAN crystal, and the cellulose I $\beta$  crystal utilized the ReaxFF parameters developed by Liu et al.<sup>17</sup> and Damirchi et al.<sup>16</sup> For iron, the ReaxFF parameters by

Islam et al.<sup>32</sup> were used, which were originally developed for analyzing hydrogen embrittlement in pure and defective iron.

*Simulation Protocol.* Stress-strain simulations under a tensile load were performed by assigning Morse parameters to the bonds that were considered the weakest according to their bond dissociation energies, for example, C-S bonds rather than C=C (double) bonds. The Morse parameters for C<sub>ar</sub>-C<sub>ar</sub> partial double bonds were used for graphite and CNTs, the Morse parameters for C-C single bonds were used for the PAN crystal, and Morse bond parameters for cellulose included C-C single-bonds and C-O single bonds. No Morse bond parameters were needed for iron since it only involves Lennard-Jones potential and no bonded terms in IFF. The inclusion of a metal remains important here to showcase how IFF-R is applicable to all types of materials and their mixtures, even when no Morse bonds are required due to the specific type of bonding.

The simulation of tensile stress-strain curves involved equilibration of the model systems using molecular dynamics simulations in the NVT ensemble at 298 K for 10 picoseconds. Following equilibration, molecular dynamics simulations were carried out in the NPT ensemble to obtain the stress-strain curves, except for the SWCNT which was run using the NVT ensemble due to inclusion of some vacuum around the CNT (the cross-sectional area pertaining to the SWCNT only was subsequently calculated). Simulations of the stress-strain curves under tensile load employed 1 atm pressure in the lateral directions and a constant rate of strain of  $2 \cdot 10^{-6}$  engineering strain every 10 timesteps of 0.5 fs along the tensile direction (0.4 per ns). Some simulations used higher strain rates of up to  $1.0 \text{ ns}^{-1}$  which did not change the results.  $\gamma$ -iron was strained in the [111] direction perpendicular to the (111) surface. The molecular dynamics simulations involved a timestep of 0.5 femtoseconds to allow small, equilibrium displacements and accurate recordings

of strain, stress, and energy. The simulations were continued until material failure, or a maximum of 100% strain.

*Analysis.* The stress was recorded as a function of strain to draw the computed stress-strain curves. The stress was obtained according to the virial theorem relative to the initial cross-sectional area of the box perpendicular to the tensile direction for all systems except the SWCNT. The model system of the SWCNT included vacuum in the simulation box, and the stress relative to the cross-sectional area of the simulation box was normalized to the cross-sectional area of the CNT only, excluding the contribution by vacuum in the simulation cell. The normalized area (in nm) was obtained using the equation  $A = N\pi(r + 0.19 \text{ nm})^2$  where  $N$  is the number of CNTs,  $r$  is the initial (equilibrium) radius of a CNT, and 0.19 nm corresponds to the addition of the van-der-Waals radius (equal to half of the layer spacing of graphite of 0.17 nm, plus a small allowance of 0.02 nm to account for the non-planar geometry and hexagonal packing of CNTs). The engineering strain was calculated as the relative expansion of the simulation cell in the tensile direction relative to the original length.

The Young's modulus was calculated by applying a line of best fit to the linear stress-strain response of the materials models near zero stress, typically between 0.00 to 0.01 tensile strain, and the slope of the best fit line was determined to be the Young's modulus ( $E = \Delta\sigma/\Delta\varepsilon$ ). The tensile strength was defined as the maximum stress at or before material failure.

The MD simulation trajectories were visualized using the OVITO v2.8 software.<sup>33</sup> Inspection of all trajectories as a function of simulation time ensured the absence of unexpected flaws in the models and associated stress-strain curves for analysis (e.g. human error through missing atoms or in setup protocol). Simulations were repeated in case issues were found. Atomic coordinates were printed to the trajectory file every 1000 timesteps. To visualize the exact instance of bond failure,

a separate simulation was carried out in which the models were strained up to approximately 2000 timesteps before failure. Afterwards, atomic coordinates were printed every 1 to 10 timesteps to capture the instance of initial bond dissociation.

**S1.3. Ab-Initio Calculation of the Young's Modulus for Polyacrylonitrile.** We utilized the CASTEP<sup>34</sup> program to compute the Young's modulus of syndiotactic, crystalline PAN in the direction of alignment of the chains as a reference relative to MD simulations with IFF-R. First, a geometry optimization was applied to an orthorhombic simulation cell with dimensions 1.00 nm x 1.15 nm x 1.16 nm and 4 aligned syndiotactic PAN chains with a density of 1.16 g/cm<sup>3</sup>.<sup>27</sup> The energy cutoff for the geometry optimization was set to 280 eV with a self-consistent field tolerance of 1.0e-6 eV/atom. The fast Fourier transform grid was set to a density of 50x72x96 with a scaling factor of 1.2 and a 1x2x1 k-point grid was utilized.

The CASTEP analysis tool in Materials Studio was then used to calculate the elastic constants of the optimized PAN structure using the GGA PBE density functional and fine quality results using a 2x2x2 Monkhorst-Pack grid of k-points. The simulation involved 4 steps for each strain with a maximum strain amplitude of 0.003. The lattice parameters were allowed to change, and convergence thresholds included an energy tolerance of 2.0e-6 eV/atom with maximum force of 0.006 eV/Å and a maximum lattice displacement of 2.0e-4 Å. Simulations were run until convergence after 120 iterations.

**S1.4. Simulation of the Surface Energy of Graphite and  $\gamma$ -Iron. *Model Building.*** An initial model of bulk graphite with 3D periodicity was created in Materials Studio with lattice parameters  $a = 1.476$  nm,  $b = 1.2874$  nm,  $c = 4.08$  nm, and  $\alpha = \beta = \gamma = 90^\circ$  (a rectangular supercell corresponding to the hexagonal unit cell). The equilibrium density was calculated to be 2.24 g/cm<sup>3</sup>, and the spacing between graphitic layers was 0.335 nm.<sup>18,19,35</sup> A model of bulk  $\gamma$ -iron was created

using the lattice parameters  $a = 1.76505$  nm,  $b = 2.1837$  nm,  $c = 1.85295$  nm, and  $\alpha = \beta = \gamma = 90^\circ$ , corresponding to a supercell with the  $z$ -axis aligned in the [111] direction and a density of 8.18 g/cm<sup>3</sup> at room temperature as given in ref. <sup>28</sup>.

Cleaved surface slabs of graphite and  $\gamma$ -iron were created after obtaining equilibrium lattice parameters for each material from an NPT ensemble simulation at 298 K and 1 atm pressure for 1 ns. The equilibrated structures were imported into Materials Studio, and the simulation cell was expanded by 4.0 nm in the  $z$ -direction, creating a solid-vacuum interface. As a result, two surfaces of equal dimension were created by introducing a vacuum slab into the 3D periodic model.

*Simulation Protocol.* First, the models were subjected to energy minimization with an energy tolerance of  $10^{-4}$  and a force tolerance of  $10^{-6}$  kcal/(mol·Å). Second, molecular dynamics simulations at 298 K and 1 atm pressure in the NPT ensemble were performed for 1 ns to obtain structures in equilibrium and average lattice parameters. Third, MD simulations were continued in the NVT ensemble at 298 K for 1 ns using the average lattice parameters and utilized to obtain the total average energy of the equilibrated bulk material and for the cleaved structure. The LJ cutoff was set to 12.0 Å (IFF standard) and electrostatic interactions were computed in high accuracy using the PPPM kspace solver with an accuracy of  $10^{-4}$  (and including a cutoff of 8.0 Å or higher for the direct summation, which has no influence on the results).

*Analysis.* The time-averaged total energy for the last portion of 1 ns in equilibrium was recorded as  $E_{\text{bulk}}$  and  $E_{\text{cleaved}}$ , and then used to calculate the surface energy of graphite and  $\gamma$ -iron, whereby  $A$  was the surface area of the slabs

$$\gamma = \frac{E_{\text{cleaved}} - E_{\text{bulk}}}{2 \times A}. \quad (\text{S1})$$

The surface energies are reported in mJ/m<sup>2</sup>.

**S1.5. Simulation of the Enthalpy of Vaporization of Propionitrile and the Enthalpy of Sublimation of  $\alpha$ -D-Glucose.** *Model Building.* The dimensions of a model of liquid propionitrile were determined using an orthogonal simulation box containing 108 CH<sub>3</sub>-CN molecules, leading to equilibrium dimensions of  $a = 2.328$  nm,  $b = 2.328$  nm, and  $c = 2.330$  nm. These dimensions reproduced propionitrile's experimentally determined density of 0.782 g/cm<sup>3</sup> within 0.5%.<sup>10</sup> Simulations of the vapor phase assumed a large orthogonal simulation cell with dimensions of  $a = b = c = 63.0$  nm containing 20 propionitrile molecules.

A model of the crystal structure of  $\alpha$ -D-glucose was created using a superstructure based on the unit cell dimensions measured by McDonald et al.<sup>36</sup> The supercell of crystalline  $\alpha$ -D-glucose contained 64 glucose molecules and had lattice parameters of  $a = 1.858$  nm,  $b = 2.530$  nm,  $c = 2.680$  nm. The simulation cell for  $\alpha$ -D-glucose in the vapor phase contained 20 glucose molecules and had lattice parameters of  $a = b = c = 100.0$  nm.

*Simulation Protocol and Analysis.* MD simulations of liquid propionitrile and crystalline glucose were carried out in the NPT ensemble for 1 ns (following longer initial equilibration). Simulations of the vapor phases were carried out for 1 ns in the NVT ensemble. The enthalpy of vaporization or sublimation enthalpy  $\Delta H_{v,s}$ , respectively, was calculated as follows:

$$\Delta H_{v,s} = \frac{\langle E_{vap} \rangle}{\# \text{ of } vap \text{ mol.}} - \frac{\langle E_{liq,cry} \rangle}{\# \text{ of } liq, cry \text{ mol.}} + RT \quad (S2)$$

$E_{vap}$  was the average total energy of the simulation system in the vapor phase,  $E_{liq,cry}$  was the average total energy of the liquid or crystal simulation system, and  $RT$  is the gas constant 8.314 J/mol multiplied by the simulation temperature. The formula includes the necessary correction of the molar vaporization energy to the molar vaporization enthalpy ( $\Delta H = \Delta U + PV$ ), using the

ideal gas law  $pV = nRT$  to approximate the volume work under constant pressure ( $n = 1$  mol).<sup>37</sup>,

38

**S1.6. Parameterization of the Morse Bond Potential for Carbon Nanotubes and Polyacrylonitrile.** The Morse bond parameter  $\alpha$  of graphite and SWCNTs for IFF-R was iteratively obtained by computing the vibration spectra (including the Raman G band), comparison to experimental data, and adjustments of  $\alpha$  for a close match ( $<20$   $\text{cm}^{-1}$  deviation). The refinements of  $\alpha$  also improved the predicted Young's Modulus and ultimate tensile strength relative to experimental data<sup>39</sup> and quantum calculations.<sup>34, 39-42</sup> The final value of  $\alpha = 2.4 \text{ \AA}^{-1}$  was chosen to reproduce the mechanical properties and the vibration spectra well.

Similarly, small adjustments of  $\alpha$  helped to closely match the Young's modulus and tensile strength of PAN in the simulations. However, the correlation of  $\alpha$  rests with the reproduction of IR and Raman spectra as mechanical properties are additionally influenced by the bond stiffness ( $\alpha$ ), other bonded, as well as nonbonded forces.

**S1.7. PCFF Parameters for Propionitrile and PCFF Parameter Improvements for Glucose.** Improvements of the base parameters in PCFF (not Morse parameters) were made for propionitrile and  $\alpha$ -D-glucose (Figure S11). The atomic charges and the nonbonded parameters of propionitrile were reevaluated using consistent dipole moments of small nitrile molecules<sup>10</sup> and a series of molecular dynamics simulations to adjust LJ parameters using the Discover program in Materials Studio.<sup>43</sup> The overall performance of the updated PCFF parameters for propionitrile is close to IFF standards. The base parameters for glucose in PCFF seriously lack structural and thermodynamic consistency and were updated for partial improvement, still far from IFF standards, since the development of basic force field parameters is not the focus of this work.



Model building of propionitrile and  $\alpha$ -D-glucose followed the protocols described in Section S1.5. The simulation of liquid and vapor phases of propionitrile was carried out using 3D boxes with 108 molecules at a temperature of 298.15 K, velocity scaling for temperature control, a time step of 0.5 fs, and a total simulation time of at least 1 ns. The simulations utilized the NPT ensemble for the liquid phase, maintaining constant pressure at 1.00 atm controlled by the Parrinello-Raman barostat. The NVT ensemble was used for the gas phase, including large distances between individual molecules of  $>2$  nm. The summation of Lennard-Jones interactions was carried out with a spherical cutoff at 12.0 Å and electrostatic interactions were computed as an Ewald summation with a high accuracy of  $10^{-4}$ .

For analysis, the average density, the average total energy for liquid propionitrile and the average energy for propionitrile in the gas phase were calculated. The enthalpy of vaporization for propionitrile was obtained using equation (S2). Iterative modifications of the atomic charge distribution of propionitrile and of the 9-6 Lennard Jones parameters were made until the liquid density and enthalpy of vaporization were within the error of experimental observations (Table 3, Figure S11, and Table S2).<sup>10,13</sup> For example, the atomic charge on the C atom “cac” next to the nitrile group had a significant effect on the vaporization energy. The default atomic charge was +0.11 e while, to reproduce the liquid density and vaporization enthalpy known from measurements, a charge of +0.30 e, which is about 3 times the default charge, was necessary (Figure S11). Likewise,  $r_{min}$  and epsilon parameters in the Lennard-Jones potential were adjusted to account for changes in atomic radii and polarizability consistent with similar chemistry covered in IFF<sup>44</sup> and other force fields (CHARMM, OPLS-AA).

Improvements of parameters for  $\alpha$ -D-glucose followed a similar approach. Since glucose is a solid at room temperature, a model of the crystal structure<sup>36</sup> rather than a liquid was used in

addition to the gas phase, and the sublimation enthalpy instead of the vaporization enthalpy. It remains challenging to reproduce the lattice parameters for  $\alpha$ -D-glucose. Minor adjustments in atomic charges and Lennard-Jones parameters were made to reproduce the density and the enthalpy of sublimation within the range of experimental error (Table S2). Hereby, atomic charges were assigned by analogy to chemically similar compounds (alcohols, ethers) and using boundary conditions such as local charge neutrality of functional groups (e.g., for the force field types  $h_o + o_h + c_1 + h_c$  in Figure S11, the net charge is zero). Improvements to IFF quality will require a review of the charge distribution, including stereoelectronic effects, and the representation in the force field along with a review and rationale for all other parameters.

**S1.8. Simulation of Epoxy-Amine Polymerization Using REACTER.** *Model Building, Creation of Template and Map Files.* A structure with bis[4-(3-aminophenoxy)phenyl] sulfone (*m*-BAPS), bisphenol A diglycidyl ether (DGEBA), the template for the product of the two monomers to create a secondary amine, and the template for the product of the secondary amine and DGEBA to create a tertiary amine were created in a singular Materials Studio file for use with REACTER.<sup>2</sup> Usage of REACTER requires that all molecules and atom types exist in the same file (i.e., product atom types that differ from reactants). Separately, individual pre-reaction molecular topologies and post-reaction topologies were created to work with the AutoMapper tool.<sup>45</sup> All structures were exported in .car/.mdf format and converted to a LAMMPS data file using msi2lmp.exe.

The REACTER map file was created by visually inspecting the pre- and post-reaction topology files in OVITO<sup>46</sup> and manually filling out the required information for map files in a text document. As such, two map files were created for the primary and secondary epoxy-amine reactions. The map files contain the “edgeIDs”, “InitiatorIDs”, and atom ID “equivalences” between the pre- and post- reaction molecule files.

*Equilibration Phase.* The equilibration phase of the simulation assumed usual settings for molecular dynamics simulations in the NPT ensemble as described earlier, including some important differences. Some additional memory, beyond that for non-reactive MD simulations, needs to be available to accommodate special bonds, bonds, and angles that may form during the reactions. The “bond\_style” used harmonic bonds and was set to “class2” since we employed IFF-R in PCFF format for this example (Morse bonds could also be used). The “improper\_style” was disabled and set to “none”, which avoided errors while using REACTER and the “fix bond/break” command. The newest release of LAMMPS (at the time of writing) has meanwhile added support for using "fix bond/break" with class 2 improper.

After these initial settings, we created coordinate files for individual monomers and replicated the monomers such that the simulation cell contained more than 30,000 atoms (IFF-R can handle  $>10^6$  atoms if needed). Next, the template files for pre- and post- reaction molecules were read and assigned specific names. The initial structure was subjected to energy minimization to relax the initial, randomly distributed, and loosely packed monomers. To create a bulk model system of equilibrium density, electrostatic calculations were turned off, the molecular model was subjected to MD simulation at 300 K and 1000 atm pressure for 50 ps to compress the reactants into a bulk structure. The electrostatic interactions were turned back on and the simulation cell was allowed to equilibrate in MD simulation at 300 K and 1 atm pressure for 50 ps. Turning off the electrostatic interactions during the compression into a bulk structure was necessary in LAMMPS, as otherwise the code may lose track of atoms (it is possible this is a glitch in LAMMPS since turning off electrostatic interactions in the interim is not necessary in other MD programs such as Discover, NAMD, or GROMACS).

*Reaction Phase.* After equilibration, the barostat was unfixed and the REACTER settings were applied by using “fix bond/react”. We used the “stabilization” option with “stabilize\_steps” set to 500 fs. The possibility for reactions was checked every 100 fs with the minimum distance set to 0 Å, the maximum distance set to 6 Å, and the probability set to 0.1. Additionally, the Nose-Hoover thermostat was applied to all atoms not participating in the reaction with a temperature of 600 K. A separate velocity rescaling thermostat was applied to all reacting atoms with temperature rescaling set to 200 K. The choice in thermostats ensures that non-reacting atoms can efficiently move into reactive configurations during the limited simulation times of 150 ps, and reacting atoms mitigate strong attractive forces upon bond connection as well as strong repulsive forces when related nonbonded atoms happen to have close contacts or overlap during the reaction. The reaction simulation was run for a total time of 49.95 ps. Hereby, the runtime was deliberately not divisible by the time interval REACTER checks for reactions (0.1 ps) to avoid a situation in which the simulation could conclude with unresolved reactions and errors before the next step. Overall, the REACTER method is a heuristic approach to model chemical reactions, and requires further development towards ergodicity.

*Strain and Failure Phase.* Bond parameters were reassigned to Morse bonds after the reaction phase of the simulation to ensure a continuous simulation as bonds form and break. The “bond\_style” was redefined to “morse” or “hybrid morse class2” before starting the “strain phase”. Thus, Morse bond parameters should be made available at the outset. The “fix bond/react” settings were unfixed in reverse order in which they are applied, which is necessary to avoid an error. Next, the “fix bond/break” options should be applied or read in from `auto_morse_bond_update.py`, which will apply the disconnection of Morse bonds past a pre-defined bond elongation, e.g., 200% (Section S2). IFF-R simulations proceeded as normal MD simulations without additional options,

however, multiple Nose-Hoover thermostats can also be applied as in the “fix bond/react” settings. The simulations of stress-strain curves up to failure were then applied as described in section S1.2.

## **S2. A Graphical User Interface and Python Code to Automatically Convert Harmonic Bonds into Morse Bonds**

A graphical user interface (GUI) and the `auto_morse_bond_update.py` code in the Supplementary Software (“`CODE_AND_GUI_FOR_AUTO_MORSE_BOND_ASSIGNMENT`”) automate the process of switching harmonic bonds to Morse bonds (Figure S1). The code provides a LAMMPS script that can be read in with the “include” statement so that “fix bond/break” commands can be quickly and easily used for each newly assigned Morse bond. The input to the code is a LAMMPS data file that has classical bond parameters. The outputs include a data file with Morse bond parameters, a LAMMPS input script file containing “fix bond/break” commands for the newly assigned Morse bonds, a pdf file showing Morse bond parameterization graphically, and a logfile with system information that can be useful for tracking parameter assignment.

To convert harmonic bonds to Morse bonds using `auto_morse_bond_update.py` (Figure S1), we use a graph theory-based approach with chemical intuition to find the bond order of each `bondTypeID` in the LAMMPS datafile. We then assign an experimental bond dissociation energy to that `bondTypeID` using the values found in our expandable database “`src/auto_morse_bond/Morse_parameters.txt`”, included in the Supplementary Software. Subsequently, the alpha parameter is fitted by minimizing the sum of residuals of the harmonic bond and the Morse bond near the equilibrium position, enabling an exact match of the Morse bond to the harmonic potential at small bond elongation. Lastly, the code writes a new LAMMPS

datafile with the new Morse bond parameters in a hybrid style. Hereby, certain bondTypeIDs may remain of the harmonic type since only weak bond types that are exposed to load and prone to disconnect need to be converted to a Morse bond. Figure S1 shows an overview of the GUI and further options.

### **S3. Molecular Response Upon Bond Dissociation and Recommended Bond Cutoffs**

**S3.1. Molecular Strain and Topology Upon Dissociation of Morse Bonds.** When using a Morse bond potential to simulate bond dissociation, the bond, angle, torsion, improper, cross terms (if present), and nonbonded interactions in the energy expression contribute to the total energy regardless of how much the bond has been elongated apart from equilibrium. Using a Morse bond alone, also does not disconnect the originally bonded atoms nor turn off the other energy contributions. The first problem can be solved using a maximum Morse bond length ( $R_{max, ij}$ ), or optionally including a probability of bond cleavage at  $R_{max, ij}$  for each ‘reactive’ bond. For example, the command “fix bond/break” in the simulation code LAMMPS can be employed to disconnect exposed bonds, turn off all other associated energy contributions (angle, dihedral, etc), and enable new nonbonded interactions between the dissociated atoms, e.g., van der Waals and Coulomb interactions between formerly 1, 2 connected and 1,3 connected atoms.

In addition, the disconnection of bonds at strains relatively far from equilibrium typically releases a large amount of strain energy in the form of increased particle velocities. The sudden release of such strain energy, which can include many neighbor atoms, can be mitigated using velocity rescaling (e.g. a temperature window of 10 K), which we have used here, and which appears to be sufficient for a smooth progression of the simulation. Alternative methods to remove

excess energy from the system include other thermostats, such as velocity rescaling through canonical sampling,<sup>47</sup> smoothly adjusting bonded interactions to zero using additional parameters, such as in the development of the AIREBO potential<sup>48</sup> and as utilized for the dissociation of polymers,<sup>49</sup> or by using REACTER with the "stabilization" option instead of "fix bond/break". However, using REACTER to break bonds will involve the creation of bond break templates which require more user effort than "fix bond/break".

**S3.2. Dissociation of Morse Bonds and Bond Cutoffs.** The dissociation of Morse bonds, for example, by using "fix bond/break" in LAMMPS, is not strictly necessary to simulate stress-strain properties of materials up to failure by bond dissociation. For a more realistic analysis at and post-failure for a chemical system, however, we propose a shifted Morse potential with a bond cutoff at about 200% of initial bond length. Systematic testing of various cutoff distances for polymers and carbon nanotubes suggests this value, which also is close to the inflection point of the Morse bond force vs bond length curve. The inflection point of a Morse bond force vs bond length curve represents the point at which the attractive force of a Morse bond is overcome by the strain energy. The bond energy at this cutoff distance then reaches zero, allowing a relatively unbiased analysis of energy changes in 3D atomic structures upon failure, especially when the strain continues far beyond the strain at break (Figure 1 and Figure S2).

The default Morse potential (left side of Figure S2a) is not suited as it introduces an energy discontinuity upon bond scission (e.g., using the "fix bond/break" command), as the bond energy then drops from  $D_{ij}$  to 0. A shifted Morse potential (right side of Figure S2a), on the contrary, minimizes discontinuities in bond energy. The implementation of the shifted Morse potential is possible, with moderate effort, in any MD program. As an example, LAMMPS requires modification of the source code (provided in the Supplementary Software as "user Morse").

Hereby, we found that the shifted Morse potential may not need to result exactly in zero energy at the cutoff distance and may be applied for a range of bond cutoff distances  $R_{max, ij}$ . For several tested examples, we obtained similar results, however, with some dependence on the bond cutoff distance (Figure 2b-e). Once  $R_{max, ij}$  is sufficiently large, the results converge (Figure S2b-e and Figures S3 to S5). The minimum bond cutoff distance depends on the type of bond, in our instances of CNTs and polyacrylonitrile, it was between 140% and 170% of the equilibrium bond length. On the other hand, bond cutoff distance is unlikely larger than 2 or 3 times the original bond distance as then the remaining localized electron density for covalent bonding goes to zero. Other remaining bonded energy terms such as angles, dihedrals, and impropers increasingly lose influence and become unphysical, and nonbonded interactions between the dissociating atoms and connected groups of atoms begin to play a role. Therefore, we recommend 200% of the original bond length as a cutoff distance, and no cutoffs larger than 300% of the original bond length.

In more detail, arbitrarily selecting a “fix bond/break” cutoff distance can cause premature failure and misleading mechanical properties (Figure S3). For example, using a single polyacrylonitrile (PAN) chain with a 1.54 Å equilibrium C-C bond distance ( $r_0$ ), we found that failure occurred at lower strain values for cutoffs less than 2.6 Å and remained unchanged for cutoff distances of 2.6 Å and higher. The total energy of the PAN tensile simulation spikes prematurely for cutoffs less than 2.6 Å and shows a large irregularity for even shorter bond cutoffs such as  $R_{max, C-C} = 1.7$  Å (Figure S3a, c). Small spikes in the total energy at the strain of bond breaking are commonly observed for any bond cutoff, whereby the total energy as a function of strain remains unchanged for cutoffs greater than 2.6 Å, equal to >170% of the original bond length. The spikes are likely the result of strain energy release from an overall strained backbone, and the simultaneous addition of nonbonded interactions between the formerly bonded atoms.



A SWCNT was used as another example for comparing the effect of bond cutoff distances on the stress-strain curves and energies (Figures S4 and S5). Early onset of failure is shown for  $R_{max, C-C}$  values less than 2.0 Å (Figure S4a) and, like PAN, premature spikes in the total energy were observed for even shorter bond cutoff distances (Figure S3c). Values of  $R_{max}$  greater than or equal to 2.0 Å show nearly identical, overlapping stress vs strain curves as well as curves of total energy vs strain (Figure S3b, d). There appears to be no upper limit for a bond cutoff up to 4.0 Å, and we recommend  $R_{max}$  to be set at ~200% of the original bond length  $r_0$ . If a shorter cutoff is desired, it should be at least 170% of the length of a C-C single bond (2.6 Å) for polymers such as PAN and at least 140% of the length of an aromatic C-C bond such as graphite.

The magnitude of the spikes in total energy at bond dissociation is susceptible to the shift of the Morse potential and the new nonbonded parameters for the reaction products (atomic charges and LJ parameters), and not significantly to the cutoff distance  $R_{max}$  in the recommended range. A reduction of the small energy spikes at bond breaking, if desired, is thus possible by modification of the shift of the Morse potential, which can be chosen different from  $D_{ij}$ . Smoothing functions with additional empirical parameters could also be employed.

**S3.3. Necessity for a Shifted Morse Potential and Analysis of Individual Energy Contributions During Failure.** The spontaneous spikes in energy upon failure, as shown for PAN (Figure S3), and SWCNTs (Figure S4) were further analyzed in terms of the individual energy contributions for the example of SWCNTs (Figure S5). Using the “fix bond/break” command ( $R_{max} = 2.84$  Å for CNT) with or without shifting the Morse potential, allows continued simulations of dissociated SWCNTs beyond the failure strain of 20% to higher strains (e.g., >40%). When not allowing for bond breaking (with or without a shifted Morse potential), the simulations become unstable and terminate once the bonds are separated by multiples of the

equilibrium bond lengths. Termination can result from atoms that participate in the bonds/angles/dihedrals involved at the site of dissociation and migration outside the range used in spatial decomposition with multi-processor computing (becoming “lost”).

Using the unshifted Morse potential, the total potential energy, bond energy, and angle energy return to a baseline value after bond dissociation after reaching the bond cutoff distance (Figure S5a-d). A spike in kinetic energy is observed upon failure at 20% strain in all instances, regardless of whether the bonds are subsequently disconnected or not, due to a gain in atom mobility (Figure S5b). Bond and angle energies are lower when permitting bond disconnection compared to not permitting bond disconnection due to the opportunity for relaxation (Figure S5c, d). Using the modified (shifted) Morse potential, the potential energy and bond energy are higher than the baseline value after dissociation without a shift of the Morse potential. The difference equals the product of the number of bonds dissociated and the bond’s dissociation energy.

The van-der-Waals energy shows a spike and then decreases (Figure S5e). Van-der-Waals interactions for severed bonds are turned on upon bond disconnection and become attractive, as is seen by the decreasing energy when compared to the examples where “fix bond/break” was not used (Figure S5e). The electrostatic energy increases (Figure S5f) when permitting bond disconnection due to switching on previously excluded electrostatic interactions (and retaining atomic charges after bond breaking for simplicity here). The higher electrostatic energy after bond dissociation for CNTs, as an example, results from the negative charge on the virtual  $\pi$ -electrons and the positive charges on the carbon atoms, which repel each other after failure. Other contributions from improper energy and cross terms (not shown in Figure S5) also increase and lead to an overall higher plateau of the total potential energy after bond dissociation when using

the modified Morse bond potential with the “fix bond/break” command than without “fix bond/break” (Figure S5a).

**S3.4. Chemical Reactions After Bond Dissociation.** From a chemistry perspective, bond scission results in the formation of radical or charged species, followed by secondary reactions depending on the composition and state variables of the system (Figure S6). Existing knowledge of reaction mechanisms and theory need to be used to guide the simulation towards realistic reaction outcomes. The newly formed chemical species need to be represented following the equation for the chemical reaction, associated stoichiometries, and by updated force field parameters. For example, new atom types and associated parameters may be assigned to the reaction products, which may often be available, or intermediate reactive chemical species can be represented by dedicated new atom types and force field parameters (radicals, radical ions, carbocations, carbanions, etc.). The latter scenario can be more challenging; however, it can be achieved using IFF protocol and high-level quantum mechanics.

For the test cases presented here, atom types and force field parameters other than charge remained the same after dissociation. It is unlikely that dissociated atoms spontaneously form bonds on-the-fly during a tensile simulation at high strain rates typical of molecular dynamics simulations. For more detailed possible reactions, some pathways are suggested in Figure S6 and can be implemented as heuristic changes, e.g., as implemented in REACTER. Upon new charge assignments, overall charge neutrality of reaction sites needs to be maintained. Otherwise, automatic protocols may unintentionally create systems with a net charge and unphysical results.

For CNTs and other graphitic materials, bond dissociation under tension is likely homolytic and two radicals will form. In this scenario, assigning new charges (zero) and adjustments to the Lennard-Jones well depth (epsilon) would seem reasonable. Follow-on reactions with H<sub>2</sub>O, O<sub>2</sub>,

N<sub>2</sub>, or other chemical species present in air may need to be defined. Heterolytic cleavage of C-O bonds is a likely scenario for bond scission in cellulose. Two new force field types would be suitable for the positively charged carbon and negatively charged oxygen atoms. Intermediate parameterizations of this kind could be avoided if the products are assumed to immediately stabilize as a sp<sup>2</sup> carbon or hydroxyl group which already have force field parameters defined in PCFF-R or other suitable force fields (Figure S6). Similarly, the failure of PAN would require product specifications or new nonbonded parameters for radicals or carbocations.

In summary, chemical reactions are diverse and follow-on reactions need to be considered case-by-case taking into account existing knowledge and available experimental data.

**S3.5. Custom LAMMPS Option for a Shifted Morse Bond Potential and IFF-R.** For easy use of IFF-R, we developed a custom LAMMPS option for IFF-R that can be installed when compiling LAMMPS (“CODE\_FOR\_CUSTOM\_LAMMPS\_OPTION\_USER\_MORSE-2Aug23” in the Supplementary Software). When installing custom options during compilation, the command

“make yes-user-morse”

should be issued. The new LAMMPS option defines a Morse potential with a shift such that the bond energy can reach 0 kcal/mol at the “fix bond/break” cutoff distance, which was not previously available and is needed for monitoring bond energy in bond breaking simulations in high accuracy and compatibility with IFF-R, IFF, CHARMM, AMBER, CVFF, PCFF, COMPASS, and other force fields for multiple materials classes as described. The syntax for the IFF-R Morse bond potential is as follows:

bond\_style morse

bond\_coeff i d0 alpha r0 [rcut [offset]]

The potential form of this equation is as follows:

$$E_{Morse} = d0 * (1 - \exp(-alpha * (r - r0)))^2 - E_{rcut}[-offset] \quad (S3)$$

In the command, *i* specifies the index of the bond type in the LAMMPS data file, *d0* the bond dissociation energy in kcal/mol, *alpha* the width of the potential, and *r0* the equilibrium distance in Å (Figure 1). *E<sub>rcut</sub>* specifies the bond energy at the bond cutoff distance *rcut*, e.g., at 200% of the original bond length, and shifts the Morse bond potential to zero at *rcut* automatically. However, *offset* enables a manual shift of the Morse potential that allows further adjustment of the bond energy towards the value of *d0* (*d0* > *E<sub>rcut</sub>*). For typical use, no offset is needed (*offset* = 0).

When *rcut* is not specified, the bond energy is not offset and the potential energy behaves similar to the original Morse bond energy potential. When *rcut* is specified but not the *offset*, the Morse potential is shifted by the energy of the bond at the bond cutoff distance (*E<sub>rcut</sub>*) which is internally determined, alleviating the user from the burden of calculating the cutoff energy and the needed offset. Otherwise, if the energy *offset* is specified by the user, the specified offset is used instead. In this manner, the user can freely customize the shift in the Morse potential energy. For example, by shifting it manually so that a bond energy of 0 kcal/mol is reached at the “fix bond/break” *rcut* distance, to the bond dissociation energy (*d0*) by specifying *d0* as the *offset*, or by anything else the user desires with proper justification.

#### **S4. Bond Order Analysis of Aromatic Amines Using Resonance Structures**

The high-level quantum mechanical bond scans (MP2) show an interesting correlation between the bond order and the dissociation energy (Figure 2). The relative dissociation energies for the

same type of bond with different substituent patterns can be qualitatively described through the resonance structures, shown here for a DETDA molecule and a 4, 4'-DDS molecule (Figure S7). For DETDA, we counted the sum of all single bond versus all double bond representations for identical bonds including molecular symmetry (1, 2 and 2, 3; 3, 4 and 1, 6; 4, 5 and 5, 6) (Figure S7a). The statistics indicates that the double bond character increases in the order  $3,4 < 1, 2 < 5, 6$ , consistent with the results by MP2 calculations, especially up to 150% of the original bond length (Figure 2a, b).

A similar analysis for 4, 4'-DDS reveals a push-pull effect of electron density between the amino groups and the sulfonate groups in para position (Figure S7b). The 2,3 carbon-carbon bonds in 4,4'-DDS have almost full double bond character and are  $\sim 30$  kcal/mol stronger than the 1,2 and 3,4 bonds (Figure 2c, d).<sup>50</sup> Stereoelectronic effects, such as double bond character and related differences in bond dissociation energies, can be incorporated into IFF-R and other force fields (PCFF-R, CHARMM-R, AMBER-R, OPLS-R) using the Morse bond parameters.

## **S5. Analysis of Mechanical Properties of Defective CNT Structures and of Spider Silk Protein (Spidroin)**

Understanding the role of defects and of order versus disorder in materials across scales on their functions is a continuing challenge in materials science. IFF-R can contribute to understanding non-linear behavior under deformation and failure. We illustrate some advanced calculations of the role of defects in CNTs (Figure S8) and of the failure of the spider silk protein spidroin (Figure S9). 4 models of a SWCNT were created with an increasing number of vacancy defects (Figure S8a). Calculations of the stress-strain curves indicate that a single vacancy reduces the tensile

strength by approximately 15% (Figure S8b). Subsequent vacancies lead to a less pronounced impact on strength. At the same time, the modulus at low strain and the initial shape of the stress-strain curve are not noticeably affected by the presence of one or multiple vacancies. Larger defects, such as Stone-Wales, broken ends in CNT bundles, and defective multilayer graphitic structures in carbon fiber can also be investigated, which is beyond the scope of this work.

At times, it is desirable to predict the temperature-dependent mechanical behavior, especially for materials subjected to extreme environments. We considered the elasticity and failure mechanism of spider silk protein (spidroin), which contains 155 amino acids per strand and is approximately 77.5 nm long (assuming each amino acid is roughly 0.5 nm in width) at two different temperatures (Figure S9). This example shows how IFF-R can be applied to biomolecular force fields such as CMARMM36, AMBER/GAFF, and biochemical problems (Figure S9). We built a model containing 2 spidroin protein strands using CHARMM36m (Figure S9a) and subjected the models to stress-strain simulations at room temperature (298 K) and at a temperature of 1 K near absolute zero for comparison (Figure S9b, c). At 298 K, entanglements of the long protein strands occurred and led to bond breaking after high elasticity and ductility (~400% engineering strain). The elastic modulus, near zero strain, was computed to be ~3.6 GPa. These simulated properties agree with experimental observations<sup>4-7</sup> and could be explored further for larger model sizes, multiple configurations, spidroin mutations and different amino acid sequences (Figure S9b). In contrast, at a temperature of only 1 K, the mobility of the atoms and of the protein strands was two orders of magnitude more restricted. Based on our start structure, entanglements did not form, and the strands slid apart without bond failure upon applied stress (Figure S9c). The tensile strength, or shear strength, was then lower at approximately 0.7 GPa (Figure S9b, c), and the elastic modulus was somewhat higher at 1 K than at 298 K.

Spidroin would likely demonstrate brittle fracture at temperatures near absolute zero, e.g., if using a model of a supercooled, entangled spidroin protein (e.g. the start structure at 298 K after equilibration). If placed under strain at 1 K, stress concentration at the entanglements would lead to failure since few bonds are initially aligned in the strain direction and the relaxation time towards alignment in response to strain would be very long at 1 K versus quite short at 298 K. However, we used aligned strands and the timescale of the simulation allowed lateral motion with an even stress concentration, also showing a dependence of the failure mechanism on the initial structure (Figure S9c).

## **S6. Comparison of Computational Speed for Simulations of Large Systems**

The speed up by IFF-R for large systems of 100,000 to 200,000 atoms relative to ReaxFF is about the same as for smaller systems with 1,000 to 5,000 atoms, i.e., 20-30x for polymers, 50x for metals, and 8x for CNT bundles (Figure S10 and Figure 4). In addition, simulations with ReaxFF are at least 10x more memory-intensive than simulations with IFF-R and could not be started with less than 12 CPUs due to insufficient memory (4.84 GB/CPU available). In comparison, all large systems could be run on 1 CPU using IFF-R, which is an added benefit for resource efficiency. For a comparison on equal footing, all large systems were run using 24 CPUs for both IFF and ReaxFF.

The IFF-R simulations would finish within 24 to 48 hours on the 24 CPUs, whereas the ReaxFF simulations could take up to a month. For many supercomputer job schedulers running a single computation for 24-48 hours is allowable while running a single job for a month is typically prohibited and not a preferred timeframe for many studies. Therefore, IFF-R can also serve



resource-limited researchers to simulate mechanical properties and failure of complex materials. Foremost, IFF-R achieves highest accuracy and interpretability among alternative MD models.

## **S7. Surface Energy and Vaporization Enthalpy of Various Materials Classes with IFF-R Versus Available ReaxFF Parameter Sets**

The computed surface energies and vaporization energies, respectively, for graphite,  $\gamma$ -iron, propionitrile, and  $\alpha$ -D-glucose using IFF-R and ReaxFF were compared relative to laboratory measurements (Table 3). For graphite, virtual  $\pi$ -electrons in IFF-R are essential to reproduce surface, wetting, and adsorption properties simultaneously due to associated multipole moments.<sup>51,52</sup> The computed surface energies or cleavage energies, respectively, agree with well-known experimental reference data within the experimental uncertainty of  $\sim 1\%$  (Table 3).<sup>18,19,53</sup> Simulations of graphite with ReaxFF, which do not include virtual  $\pi$ -electrons, lead to +30% deviation in cleavage energy utilizing the Liu et al. (ref. <sup>17</sup>) parameter set.

For the (111)  $\gamma$ -iron surface, both the IFF-R and ReaxFF parameters reproduce experimental values of the surface energy well with 1% and 6% deviation, respectively.<sup>54-56</sup> However, the ReaxFF parameters used in this simulation (Shin et al.<sup>57</sup>) are different from those used in the stress-strain simulation (Aryanpour et al.<sup>58</sup>) to achieve a better fit (Figure 3h and Table 1). The Aryanpour ReaxFF parameters (ref. <sup>58</sup>) would result in greater than 200% deviation of the computed surface energy from experimental reference data when used. The differences demonstrate that using the same ReaxFF parameter set for iron, deviations on the order of 100% in either surface properties or mechanical properties are inevitable, which is not the case with IFF and IFF-R.

For propionitrile, the IFF-R parameters reproduce experimental data for the vaporization energy, and the modified PCFF-R parameters replicate the enthalpy of sublimation of glucose (Table 3)<sup>20,59,60</sup> while needing major improvements to reproduce the crystal structure (Section S1.7). ReaxFF parameters perform well for glucose but not for propionitrile, which shows 30% deviation in vaporization energy (Table 3). We tested both the Liu<sup>17</sup> (Table 3) and Damirchi<sup>16</sup> ReaxFF parameter sets (Table S2). The Liu parameters provide better nonbonded properties and can be used in calculations with the *lg* van-der-Waals correction. The *lg* van-der-Waals correction is an additional vdW interaction more adequate to capture long-range dispersion forces and to reproduce mass densities in molecular crystals than the standard Morse potential in typical ReaxFF potentials. The Liu parameters were also optimized using the enthalpy of sublimation of solids and appear to be overall a better fit to compare surface energy and vaporization enthalpies to IFF-R results. Damirchi parameters for ReaxFF lead to higher combined deviations of 11% for propionitrile and 32% for glucose (Table S2).

In more detail, ReaxFF exhibits deviations in the surface energy of graphite and in the vaporization energy of organic liquids of up to 30% using two different parameter sets by Liu and Damirchi (Table S2). The errors are similar overall yet have a different distribution. The parameter set by Liu<sup>17</sup> includes long range van-der-Waals corrections, which may be considered an improvement over the parameter set by Damirchi,<sup>16</sup> although not showing lower errors. Improvements of ReaxFF to reduce uncertainties of 10% to 30% to IFF and IFF-R standards of ~5% would require a time-consuming reparameterization or may not be possible without enlarging errors for other chemistries covered by the same parameters.

The Damirchi parameters of ReaxFF<sup>16</sup> overpredict the surface energy of graphite with about 11% deviation from experimental data, which is an improvement over the Liu parameters<sup>17</sup> with

30% deviation. The improvement originates from optimization for flattened carbon nanotubes, informed by IFF-PCFF,<sup>16,61</sup> that are similar in structure and composition to graphitic layers. The Damirchi parameters also predict the enthalpy of vaporization of propionitrile with about 11% deviation from the experimentally determined value, which is better than 36% deviation with the vdW-corrected Liu parameters. Conversely, the Liu parameters for ReaxFF predict the enthalpy of sublimation of  $\alpha$ -D-glucose with only 4% deviation, which is much better than the Damirchi parameters with more than 30% deviation from experiment. The choice of ReaxFF parameter set is therefore often a trade-off that involves significant errors somewhere. In comparison, deviations with IFF-R are smaller and remain below a maximum of 3% (Table 3).

### Supplementary References

- 1 Stephenson, G. R. Organic chemistry. *Chem. Ind.* **85**, 42-43 (2021).
- 2 Gissinger, J. R., Jensen, B. D. & Wise, K. E. Reactor: A heuristic method for reactive molecular dynamics. *Macromolecules* **53**, 9953-9961 (2020).
- 3 Huang, J. et al. Charmm36m: An improved force field for folded and intrinsically disordered proteins. *Nature methods* **14**, 71-73 (2017).
- 4 Du, N. et al. Design of superior spider silk: From nanostructure to mechanical properties. *Biophys. J.* **91**, 4528-4535 (2006).
- 5 Griffiths, J. R. & Salantri, V. R. The strength of spider silk. *J. Mater. Sci.* **15**, 491-496 (1980).
- 6 Rising, A. et al. Spider silk proteins -: Mechanical property and gene sequence. *Zool. Sci.* **22**, 273-281 (2005).
- 7 Yarger, J. L., Cherry, B. R. & van der Vaart, A. Uncovering the structure-function relationship in spider silk. *Nat. Rev. Mater.* **3**, 8000 (2018).
- 8 Xu, Y. & Huang, J. Validating the charmm36m protein force field with lj-pme reveals altered hydrogen bonding dynamics under elevated pressures. *Commun. Chem.* **4** (2021).
- 9 Nishiyama, Y., Langan, P. & Chanzy, H. Crystal structure and hydrogen-bonding system in cellulose 1 beta from synchrotron x-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **124**, 9074-9082 (2002).
- 10 Lide, D. R. *CRC handbook of chemistry and physics* 96th edn, (CRC Press, Boca Raton, FL, 2015).
- 11 Blanksby, S. J. & Ellison, G. B. Bond dissociation energies of organic molecules. *Acc. Chem. Res.* **36**, 255-263 (2003).
- 12 Darwent, B. D. in *Nat. Stand. Ref. Data Ser., Nat. Bur. Stand. (U. S.)* Vol. 31 (ed U.S. Department of Commerce) 9-47 (Washington, D.C., 1970).

- 13 *NIST chemistry webbook*. (US Department of Commerce, DOI: doi.org/10.18434/T4D303, 2018).
- 14 Zavitsas, A. A. The relation between bond lengths and dissociation energies of carbon-carbon bonds. *J. Phys. Chem. A* **107**, 897-898 (2003).
- 15 Zeinalipour-Yazdi, C. D., Loizidou, E. Z. & Chutia, A. Size-dependent bond dissociation enthalpies in single-walled carbon nanotubes. *Chem. Phys. Lett.* **731**, 136628 (2019).
- 16 Damirchi, B. et al. Reaxff reactive force field study of polymerization of a polymer matrix in a carbon nanotube-composite system. *J. Phys. Chem. C* **124**, 20488-20497 (2020).
- 17 Liu, L. C., Liu, Y., Zybin, S. V., Sun, H. & Goddard, W. A. Reaxff-g: Correction of the reaxff reactive force field for london dispersion, with applications to the equations of state for energetic materials. *J. Phys. Chem. A* **115**, 11016-11022 (2011).
- 18 Liu, Z. et al. Interlayer binding energy of graphite: A mesoscopic determination from deformation. *Phys. Rev. B* **85**, 205418 (2012).
- 19 Zacharia, R., Ulbricht, H. & Hertel, T. Interlayer cohesive energy of graphite from thermal desorption of polyaromatic hydrocarbons. *Phys. Rev. B* **69**, 155406 (2004).
- 20 Oja, V. & Suuberg, E. M. Vapor pressures and enthalpies of sublimation of d-glucose, d-xylose, cellobiose, and levoglucosan. *J. Chem. Eng. Data* **44**, 26-29 (1999).
- 21 Pople, J. A., Binkley, J. S. & Seeger, R. Theoretical models incorporating electron correlation. *Int. J. Quantum Chem.* **10**, 1-19 (1976).
- 22 Becke, A. D. A new mixing of hartree-fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372-1377 (1993).
- 23 Gaussian 16 rev. C.01 (Wallingford, CT, 2016).
- 24 Singh, S. K. et al. Thermal properties of fluorinated graphene. *Phys. Rev. B* **87**, 104114 (2013).
- 25 Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1-19 (1995).
- 26 Pramanik, C., Gissinger, J. R., Kumar, S. & Heinz, H. Carbon nanotube dispersion in solvents and polymer solutions: Mechanisms, assembly, and preferences. *ACS Nano* **11**, 12805-12816 (2017).
- 27 Qin, Y. in *Med. Text. Mater.* (ed Yimin Qin) Ch. 3, 23-42 (Woodhead Publishing, 2016).
- 28 Kanhaiya, K., Kim, S., Im, W. & Heinz, H. Accurate simulation of surfaces and interfaces of ten FCC metals and steel using lennard-jones potentials. *npj Comput. Mater.* **7**, 17 (2021).
- 29 Basinski, Z. S., Humerothery, W. & Sutton, A. L. The lattice expansion of iron. *Proc. R. Soc. London, Ser. A* **229**, 459-467 (1955).
- 30 Sun, H., Mumby, S. J., Maple, J. R. & Hagler, A. T. An ab-initio cff93 all-atom force-field for polycarbonates. *J. Am. Chem. Soc.* **116**, 2978-2987 (1994).
- 31 Sun, H. 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).
- 32 Islam, M. M., Zou, C., van Duin, A. C. T. & Raman, S. Interactions of hydrogen with the iron and iron carbide interfaces: A reaxff molecular dynamics study. *Phys. Chem. Chem. Phys.* **18**, 761-771 (2016).
- 33 Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO-the open visualization tool. *Modell. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).

- 34 Clark, S. J. et al. First principles methods using CASTEP. *Z. Kristallog.* **220**, 567-570 (2005).
- 35 Trucano, P. & Chen, R. Structure of graphite by neutron-diffraction. *Nature* **258**, 136-137 (1975).
- 36 McDonald, T. R. R. & Beevers, C. A. The crystal and molecular structure of alpha-glucose. *Acta Crystallogr.* **5**, 654-659 (1952).
- 37 Mahoney, M. W. & Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **112**, 8910-8922 (2000).
- 38 Kelkar, M. S. & Maginn, E. J. Calculating the enthalpy of vaporization for ionic liquid clusters. *J. Phys. Chem. B* **111**, 9424-9427 (2007).
- 39 Bai, Y. X. et al. Carbon nanotube bundles with tensile strength over 80 GPa. *Nat. Nanotechnol.* **13**, 589+ (2018).
- 40 Zhang, R. F. et al. Superstrong ultra long carbon nanotubes for mechanical energy storage. *Adv. Mater.* **23**, 3387+ (2011).
- 41 Zhang, R. F., Zhang, Y. Y. & Wei, F. Controlled synthesis of ultralong carbon nanotubes with perfect structures and extraordinary properties. *Acc. Chem. Res.* **50**, 179-189 (2017).
- 42 Zhao, Q. Z., Nardelli, M. B. & Bernholc, J. Ultimate strength of carbon nanotubes: A theoretical study. *Phys. Rev. B* **65**, 144105 (2002).
- 43 *Materials studio 2019 program suite and user guide.* (Biovia/Dassault Systemes, Cambridge, UK, 2019).
- 44 Heinz, H., Lin, T.-J., Mishra, R. K. & Emami, F. S. Thermodynamically consistent force fields for the assembly of inorganic, organic, and biological nanostructures: The interface force field. *Langmuir* **29**, 1754-1765 (2013).
- 45 Bone, M. A., Howlin, B. J., Hamerton, I. & Macquart, T. Automapper: A Python tool for accelerating the polymer bonding workflow in LAMMPS. *Comp. Mater. Sci.* **205**, 111204 (2022).
- 46 Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool. *Model. Simul. Mater. Sc.* **18**, 015012 (2010).
- 47 Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
- 48 Stuart, S. J., Tutein, A. B. & Harrison, J. A. A reactive potential for hydrocarbons with intermolecular interactions. *J. Chem. Phys.* **112**, 6472-6486 (2000).
- 49 Yamada, N., Oya, Y., Kato, N., Mori, K. & Koyanagi, J. A molecular dynamics simulation for thermal activation process in covalent bond dissociation of a crosslinked thermosetting polymer. *Molecules* **28**, 2736 (2023).
- 50 Smith, M. B. *March's advanced organic chemistry: Reactions, mechanism, and structure.* 7th edn, (Wiley, Hoboken, NJ, 2013).
- 51 Pramanik, C., Gissinger, J. R., Kumar, S. & Heinz, H. Carbon nanotube dispersion in solvents and polymer solutions: Mechanisms, assembly, and preferences. *ACS Nano* **11**, 12805-12816 (2017).
- 52 Akkineni, S. et al. Amyloid-like amelogenin nanoribbons template mineralization via a low-energy interface of ion binding sites. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2106965119 (2022).
- 53 Shih, C. J., Strano, M. S. & Blankschtein, D. Wetting translucency of graphene. *Nat. Mater.* **12**, 866-869 (2013).

- 54 Tyson, W. R. & Miller, W. A. Surface free-energies of solid metals - estimation from liquid surface-tension measurements. *Surf. Sci.* **62**, 267-276 (1977).
- 55 Roth, T. A. Determination of the surface and grain-boundary free-energies of titanium-alloys having a bamboo structure. *J. Met.* **37**, A65-A65 (1985).
- 56 Spencer, M. J. S., Hung, A., Snook, I. K. & Yarovsky, I. Density functional theory study of the relaxation and energy of iron surfaces. *Surf. Sci.* **513**, 389-398 (2002).
- 57 Shin, Y. K., Gai, L. L., Raman, S. & van Duin, A. C. T. Development of a reaxff reactive force field for the Pt-Ni alloy catalyst. *J. Phys. Chem. A* **120**, 8044-8055 (2016).
- 58 Aryanpour, M., van Duin, A. C. T. & Kubicki, J. D. Development of a reactive force field for iron-oxyhydroxide systems. *J. Phys. Chem. A* **114**, 6298-6307 (2010).
- 59 Linstrom, P. J. & Mallard, W. G. The NIST chemistry webbook: A chemical data resource on the internet. *J. Chem. Eng. Data* **46**, 1059-1063 (2001).
- 60 Weber, L. A. & Kilpatrick, J. E. Entropy and related thermodynamic properties of propionitrile. *J. Chem. Phys.* **36**, 829-834 (1962).
- 61 Dharmawardhana, C. C. et al. Reliable computational design of biological-inorganic materials to the large nanometer scale using interface-ff. *Mol. Simul.* **43**, 1394-1405 (2017).