Table 1: Small molecule building blocks. CG particle type, corresponding chemical building block, and examples of molecules in which such a block appears. The atoms the CG block is taken to represent are also shown in the 2D chemical structures, with T-beads, S-beads, and R-beads depicted in blue, green, and red, respectively.

<sup>\*</sup> Indicates molecules considered in this work (DOI: 10.33774/chemrxiv-2021-1qmq9); others are from the Martini 3 main paper (DOI: 10.1038/s41592-021-01098-3).

type	chemical building block	examples	
VI	C	2D	name (mapping)
P2	CH <sub>3</sub> CH <sub>2</sub> -COOH	O OH	propanoic acid (P2)
SP2	$\mathrm{CH_{3}}\mathrm{-COOH}$	HO	acetic acid (SP2)
P1	$(\mathrm{CH_3})_2\mathrm{CH}\mathrm{-OH}$	<b>OH</b>	isopropanol (P1)
TP1d	$-\mathrm{CH_2OH}$ (bonded to $-\mathrm{CF_3}$ )	OH F F	2,2,2-trifluoroethanol (SX4e-TP1d)
SN6	$-\mathrm{CH}^\dagger {=} \mathrm{C}(-\mathrm{OH}) {-} \mathrm{CH}^\dagger {=}$		*phenol ((TC5) <sub>2</sub> -SN6)
	$-\mathrm{CH_2}^\dagger - \mathrm{CH}(-\mathrm{OH}) - \mathrm{CH_2}^\dagger -$	OH OH	*cyclopentanol (SN6-SC3)
SN6d	$-\mathrm{CH}^{\dagger}\mathrm{=}\mathrm{C}(-\mathrm{NH}_{2})\mathrm{-}\mathrm{CH}^{\dagger}\mathrm{=}$	ZH2	*aniline ((TC5) <sub>2</sub> -SN6d)
TN6d	$=\!\!\mathrm{CH}^{\dagger}\!-\!\mathrm{NH}\!-\!\mathrm{CH}^{\dagger}\!=$	H	*pyrrole ((TC5) <sub>2</sub> -TN6d)
TN6	=C(-OH)-	OH	* $p$ -cresol ((TC5) <sub>2</sub> -TC5-TN6)
N6a	$-(\mathrm{CH_2})_2 - \mathrm{CH}(=\mathrm{O})$	OH)	heptanal (C1-N6a)
TN6a	-CH=N-		*pyridine ((TC5) <sub>2</sub> -TN6a)
SN5a	$-O-CH_2-O-$		*1,3-dioxolane (SN5a-TC3)
TN5a	=C $=$ O		*benzoquinone (TN5a-(TC5) <sub>2</sub> -TN5a)
N4a	$-C(=O)-O-CH_3$		*methyl-benzoate (N4a-(TC5) $_3$ )

<sup>&</sup>lt;sup>†</sup> Indicates that the (group of) atom(s) is shared with neighboring beads.

$$CH_3-C(=O)-O-CH_3 \qquad \qquad \text{methyl-acetate (N4a)}$$

$$SN4a \quad -CH_2-O-CH_2- \qquad \qquad \text{*tetrahydropyran (SC3-SN4a)}$$

$$-C(=O)-CH_3 \qquad \qquad \text{*acetophenone (SN4a-(TC5)_3)}$$

$$TN4a \quad -CH_2^{\dagger}-O-CH_2^{\dagger}- \qquad \qquad \text{*tetrahydrofuran (SC3-TN4a)}$$

$$-C\equiv N \qquad \qquad \text{*benzonitrile ((TC5)_3-TN4a)}$$

$$-CH(=O) \qquad \qquad \text{*benzaldehyde ((TC5)_3-TN4a)}$$

$$SN3a \quad -CH_2-O-CH_2- \qquad \qquad \text{*introbenzene ((TC5)_3-SN3a)}$$

$$SN3a \quad -NO_2 \qquad \qquad \text{*nitrobenzene ((TC5)_3-SN3a)}$$

$$SN2a \quad =C(-O-CH_3)- \qquad \qquad \text{*o-methylanisole (SN2a-TC4-(TC5)_2)}$$

$$TN2a \quad =CH^{\dagger}-O-CH^{\dagger}= \qquad \text{*furan ((TC5)_2-TN2a)}$$

$$-O-CH_3 \qquad \qquad \text{*methoxybenzene ((TC5)_3-SN1)}$$

$$SN1 \quad -N(-CH_3)_2 \qquad \qquad \text{*N,N-dimethylaniline ((TC5)_3-SN1)}$$

$$TN1 \quad -N(-CH_3)- \qquad \qquad \text{*introbenzene ((TC5)_3-SN1)}$$

$$SC6 \quad -CH^{\dagger}=C(-SH)-CH^{\dagger}= \qquad \text{*thiophenol ((TC5)_2-SC6)}$$

$$-CH_2-S-CH_2- \qquad \qquad \text{*tetrahydrothiophene (SC6-TC3)}$$

$$TC6 \quad =CH^{\dagger}-S-CH^{\dagger}= \qquad \text{*thiophene ((TC5)_2-TC6)}$$

$$TC5 \quad -CH=CH- \qquad \text{*benzene ((TC5)_4)}$$

$$TC5e (-)_2C=C(-)_2$$

\*naphthalene 
$$((TC5)_2$$
- $TC5e$ - $(TC5)_2)$ 

$$C4$$
  $CH_3-CH=CH-CH_3$ 



$$SC4$$
  $-CH^{\dagger}=C(-CH_3)-CH^{\dagger}=$ 



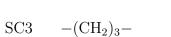
$$TC4 = C(-CH_3)-$$



\*
$$o$$
-methylanisole (SN2a-TC4-(TC5)<sub>2</sub>)

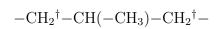
$$SX4e - CF_3$$
 $-CF_3$ 



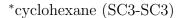




$$2,2,2$$
-trifluoroethanol (SX4e-TP1d)







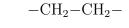


\*methylcyclopentane (SC3-SC3)

TC3



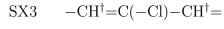
\*ethylbenzene ((TC5)<sub>3</sub>-TC3)



 $-CH_2-CH_3$ 



\*tetrahydrothiophene (SC6-TC3)





\*chlorobenzene ((TC5)<sub>2</sub>-SX3)



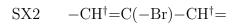


\*3-propyl-thiophene (TC6-(TC5)<sub>2</sub>-SC2)

X2  $CHCl_3$ 



\*trichloromethane (X2)





\*bromobenzene ((TC5)<sub>2</sub>-SX2)

C1 $CH_3-(CH_2)_2-CH_3$ 



butane (C1)

 $-\mathrm{CH}^{\dagger} = \mathrm{C}(-\mathrm{I}) - \mathrm{CH}^{\dagger} =$ X1

\*iodobenzene ((TC5)<sub>2</sub>-X1)