Supplementary Material

A. Proof of Proposition 1

Let $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{|V|}) \in \mathbb{R}^{|V|\times 3}$ denote a molecular conformation. Let $\gamma_{\mathbf{t}} : \mathbb{R}^{|V|\times 3} \to \mathbb{R}^{|V|\times 3}$ denote any 3D translation function where $\gamma_{\mathbf{t}}(\mathbf{R})_i \coloneqq \mathbf{r}_i + \mathbf{t}$, and let $\rho_D : \mathbb{R}^{|V|\times 3} \to \mathbb{R}^{|V|\times 3}$ denote any 3D rotation function whose rotation matrix representation is $D \in \mathbb{R}^{3\times 3}$, i.e., $\rho_D(\mathbf{R})_i = D\mathbf{r_i} \in \mathbb{R}^3$. In our context, we say a score function of atomic coordinates $\mathbf{s} : \mathbb{R}^{|V|\times 3} \to \mathbb{R}^{|V|\times 3}$ is roto-translation equivariant if it satisfies:

$$\mathbf{s} \circ \gamma_{\mathbf{t}} \circ \rho_D(\mathbf{R}) = \rho_D \circ \mathbf{s}(\mathbf{R}), \tag{11}$$

Intuitively, the above equation says that applying any translation $\gamma_{\mathbf{t}}$ to the input has no effect on the output, and applying any rotation ρ_D to the input has the same effect as applying it to the output, i.e., the gradients rotate together with the molecule system and are invariant under translation.

Proof. Denote $\hat{\mathbf{R}} = \gamma_{\mathbf{t}} \circ \rho_D(\mathbf{R})$. According to the definition of translation and rotation function, we have $\hat{\mathbf{R}}_i = \hat{\mathbf{r}}_i = D\mathbf{r}_i + \mathbf{t}$. According to Eq. 3, we have

$$\begin{aligned}
\forall i, \big(\mathbf{s} \circ \gamma_{\mathbf{t}} \circ \rho_D(\mathbf{R})\big)_i &= \mathbf{s}(\hat{\mathbf{R}})_i \\
&= \sum_{j\in N(i)} \frac{1}{\hat{d}_{ij}} \cdot \mathbf{s}(\hat{\mathbf{d}})_{ij} \cdot (\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j) \\
&= \sum_{j\in N(i)} \frac{1}{d_{ij}} \cdot \mathbf{s}(\mathbf{d})_{ij} \cdot \big((D\mathbf{r}_i + \mathbf{t}) - (D\mathbf{r}_j + \mathbf{t})\big) \\
&= \sum_{j\in N(i)} \frac{1}{d_{ij}} \cdot \mathbf{s}(\mathbf{d})_{ij} \cdot D(\mathbf{r}_i - \mathbf{r}_j) \\
&= D\Big(\sum_{j\in N(i)} \frac{1}{d_{ij}} \cdot \mathbf{s}(\mathbf{d})_{ij} \cdot (\mathbf{r}_i - \mathbf{r}_j)\Big) \\
&= D\mathbf{s}(\mathbf{R})_i \\
&= \big(\rho_D \circ \mathbf{s}(\mathbf{R})\big)_i.
\end{aligned} \tag{12}$$

Here $d_{ij} = \hat{d}_{ij}$ and $\mathbf{d} = \hat{\mathbf{d}}$ because rotation and translation will not change interatomic distances. Together with Eq. 11 and Eq. 12, we conclude that the score network defined in Eq. 3 is roto-translation equivariant. □

B. Additional Hyperparameters

We summarize additional hyperparameters of our ConfGF in Table 6, including the biggest noise level σ_1, the smallest noise level σ_L, the number of noise levels L, the number of sampling steps per noise level T, the smallest step size ϵ, the batch size and the training epochs.

Table 6. Additional hyperparameters of our ConfGF.

Dataset	σ_1	σ_L	L	T	ϵ	Batch size	Training epochs
GEOM-QM9	10	0.01	50	100	2.4e-6	128	200
GEOM-Drugs	10	0.01	50	100	2.4e-6	128	200
ISO17	3	0.1	30	100	2.0e-4	128	100

C. Additional Experiments

We present more results of the **Cov**erage (COV) score at different threshold δ for GEOM-QM9 and GEOM-Drugs datasets in Tables 7 and 8 respectively. Results in Tables 7 and 8 indicate that the proposed ConfGF consistently outperforms previous state-of-the-art baselines, including GraphDG and CGCF. In addition, we also report the **Mis**match (MIS) score defined as follows:

$$\mathrm{MIS}(S_g, S_r) = \frac{1}{|S_g|}\left|\left\{\mathbf{R} \in S_g | \,\mathrm{RMSD}(\mathbf{R}, \hat{\mathbf{R}}) > \delta, \forall \hat{\mathbf{R}} \in S_r\right\}\right|, \tag{13}$$

where δ is the threshold. The MIS score measures the fraction of generated conformations that are not matched by any ground-truth conformation in the reference set given a threshold δ. A lower MIS score indicates less bad samples and better generation quality. As shown in Tables 7 and 8, the MIS scores of ConfGF are consistently lower than the competing baselines, which demonstrates that our ConfGF generates less invalid conformations compared with the existing models.

Table 7. COV and MIS scores of different approaches on GEOM-QM9 dataset at different threshold δ.

QM9	Mean COV (%)			Median COV (%)			Mean MIS (%)			Median MIS (%)		
δ (Å)	GraphDG	CGCF	**ConfGF**	GraphDG	CGCF	**ConfGF**	GraphDG	CGCF	**ConfGF**	GraphDG	CGCF	**ConfGF**
0.10	1.03	0.14	17.59	0.00	0.00	10.29	99.70	99.93	92.30	100.00	100.00	96.23
0.20	13.05	10.97	43.60	2.99	3.95	37.92	96.12	96.91	81.67	99.04	99.04	85.65
0.30	32.26	31.02	61.94	18.81	22.94	59.66	87.15	89.21	73.06	94.44	94.44	77.27
0.40	53.53	53.65	75.45	50.00	52.63	80.64	72.60	78.35	65.38	82.63	82.63	70.00
0.50	73.33	78.05	88.49	84.21	82.48	94.13	56.09	63.51	53.56	64.66	64.66	56.59
0.60	88.24	94.85	97.71	98.83	98.79	100.00	40.36	44.82	34.78	43.73	43.73	35.86
0.70	95.93	99.05	99.52	100.00	100.00	100.00	27.93	29.64	21.00	23.38	23.38	15.64
0.80	98.70	99.47	99.68	100.00	100.00	100.00	19.15	20.98	12.86	10.72	10.72	5.23
0.90	99.33	99.50	99.77	100.00	100.00	100.00	12.76	16.74	8.98	3.65	3.65	1.53
1.00	99.48	99.50	99.86	100.00	100.00	100.00	8.00	14.19	6.76	0.47	0.47	0.36
1.10	99.51	99.51	99.91	100.00	100.00	100.00	4.99	12.26	5.57	0.00	0.00	0.00
1.20	99.51	99.51	99.94	100.00	100.00	100.00	2.95	9.68	3.48	0.00	0.00	0.00
1.30	99.51	99.51	99.96	100.00	100.00	100.00	1.65	7.48	1.91	0.00	0.00	0.00
1.40	99.51	99.51	99.96	100.00	100.00	100.00	0.84	5.94	1.05	0.00	0.00	0.00
1.50	99.52	99.51	99.97	100.00	100.00	100.00	0.41	4.94	0.70	0.00	0.00	0.00

Table 8. COV and MIS scores of different approaches on GEOM-Drugs dataset at different threshold δ.

Drugs	Mean COV (%)			Median COV (%)			Mean MIS (%)			Median MIS (%)		
δ (Å)	GraphDG	CGCF	**ConfGF**	GraphDG	CGCF	**ConfGF**	GraphDG	CGCF	**ConfGF**	GraphDG	CGCF	**ConfGF**
0.25	0.00	0.06	0.17	0.00	0.00	0.00	100.00	99.99	99.97	100.00	100.00	100.00
0.50	0.26	0.80	1.15	0.00	0.00	0.00	99.95	99.80	99.52	100.00	100.00	100.00
0.75	0.75	5.81	9.15	0.00	0.00	0.50	99.69	97.86	96.94	100.00	100.00	99.75
1.00	2.39	24.67	30.60	0.00	11.81	18.89	99.14	90.82	89.63	100.00	96.50	95.58
1.25	8.27	53.96	62.15	0.00	57.06	70.93	97.92	78.32	76.58	100.00	86.28	84.48
1.50	19.96	79.37	86.62	4.00	92.46	98.79	94.40	63.80	60.06	99.14	66.39	63.81
1.75	36.86	91.47	96.53	26.58	100.00	100.00	87.68	49.72	43.63	95.83	47.09	41.72
2.00	55.79	96.73	98.62	55.26	100.00	100.00	76.99	37.53	29.80	87.35	30.90	22.44
2.25	71.43	99.05	99.83	80.00	100.00	100.00	61.76	27.30	18.68	69.74	20.07	10.93
2.50	83.53	99.47	100.00	95.45	100.00	100.00	44.32	18.97	11.09	42.96	12.33	3.31
2.75	91.09	99.60	100.00	100.00	100.00	100.00	27.92	12.52	6.32	16.67	6.82	0.74
3.00	95.00	99.96	100.00	100.00	100.00	100.00	15.97	7.67	3.36	2.46	3.32	0.00

D. More Generated Samples

We present more visualizations of generated conformations from our ConfGF in Figure 4, including models trained on GEOM-QM9 and GEOM-drugs datasets.

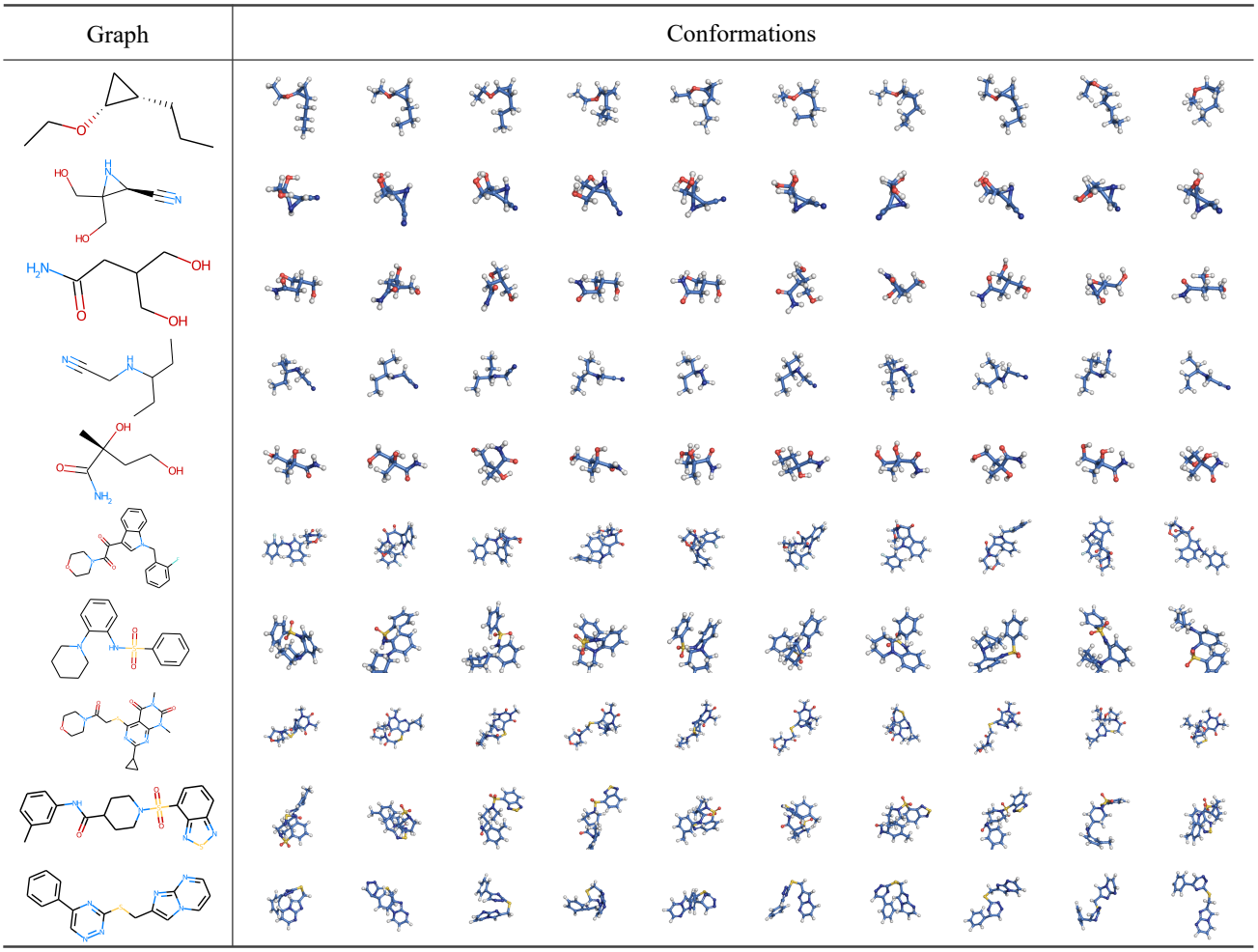

Figure 4. Visualization of conformations generated by our ConfGF. For each molecular graph, we randomly sample 10 generated conformations from our model. For the top 5 rows, the molecular graphs are drawn from the GEOM-QM9 test dataset, while for the bottom 5 rows, the molecular graphs are drawn from the GEOM-drugs test dataset.