References	Computational Models	Characteristics	*In silico* Validation	*In vivo/in vitro* Validation
[41]	SMILES based RNN-LSTM	(i) generation of compound libraries for high-throughput screening. (ii) hit to- lead optimization for targets (iii) fragment-based drug discovery.	Fine tuning against TRPM8 inhibitors	Not available
[44]	SMILES based RNN-LSTM	Model fine-tuned by transfer learning to enable the de novo generation of target-specific ligands.	Fine tuning of the model to generate molecules with agonistic activity on RXR and/orPPAR.	5 generated molecules tested in hybrid reporter gene assays for their agonistic effects on nuclear receptors RXR and PPAR *in vivo*.
[46]	SMILES based CVAE	Generate molecules with specific values for five properties (MW, LogP, HBD, HBA, and TPSA)	Applied to Aspirin and Tamiflu. property values of generated molecules within an error range of 10%	Not available
[50]	SMILES based AAE and VAE	Adversarial autoencoder applied to inverse QSAR to generate chemical structures	Novel compounds with predicted activity against dopamine receptor type 2	Not available
[51]	SMILES based ECAAE	improved model with two different disentanglement approaches and a semisupervised extension based on supervised adversarial autoencoders (SAAE)	Generated molecules evaluated using logP and SAS	Molecules tested *in vitro* for activity and selectivity against JAK3 (IC50 = 6.73 μM), while being inactive for JAK2 (IC50 = 17.58 mM), B-Raf (IC50 = 85.55 μM), and c-Raf (IC50 = 64.86 μM).
[13]	SMILES based Stack-augmented recurrent neural network	Two DNNs are trained separately and used to generate chemical libraries	99.5% of de generated JAK2 inhibitors had SAS values below 6	Not available
[59]	SMILES based GAN-RL (RANC)	RANC uses a differentiable neural computer to increase the generation capabilities and mitigate common problems found in adversarial settings.	Generated structures match the distributions of chemical features/descriptors.	Not available
[60]	SMILES based GAN-RL (ATNC)	The model uses a new objective reward function named Internal Diversity Clustering to generate more diverse molecules.	ATNC generates 72% of valid and 77% of unique SMILES. druglikeness properties estimated using chemical descriptors.	inhibition potency of generated compounds against selected kinases tested *in vitro*
[52]	SMILES based GENTRL	GENTRL prioritizes the structures it generates by using these three self-	Design novel compounds that are active against DDR1	Testing *in vitro* inhibitory activity in enzymatic kinase assay

		organizing maps (SOMs) in sequence as reward functions: the trending SOM, the general kinase SOM, and the specific kinase SOM.	kinase	for 6 compounds. One compound tested *in vivo*

Table 2. List of recent examples of applications of AI methods for de novo molecular generation. These works, referenced within the main text, have been published in peer-reviewed scientific journals. For each case study, the type of computational model and its characteristics are summarized. The performances of the models are assessed by evaluating different types of metrics. Specific metrics are used to measure the performance of the model itself in terms of overall numeric performance, convergence rate or stability. Additional metrics are used to assess the generated output. For experiments aiming at designing set of molecules with generic drug-like properties, this can include computing similarities between the generated SMILES and SMILES form the testing set, estimating the diversity and novelty of the generated set of molecules, computing chemical features and descriptors of the generated molecules. For some experiments, the model is calibrated and fine-tuned to generate molecules with good activity against a specific target. For some case studies, *in vitro* and *in vivo* validations of the most promising generated compounds were performed. Those cases are of particular interest because they illustrate the different steps (filters, synthesis, testing) each molecular structure needs to go through after the *de novo* generation procedure itself.