Supporting Information

Identification of Bioisosteric Substituents by Deep Neural Network

Peter Ertl

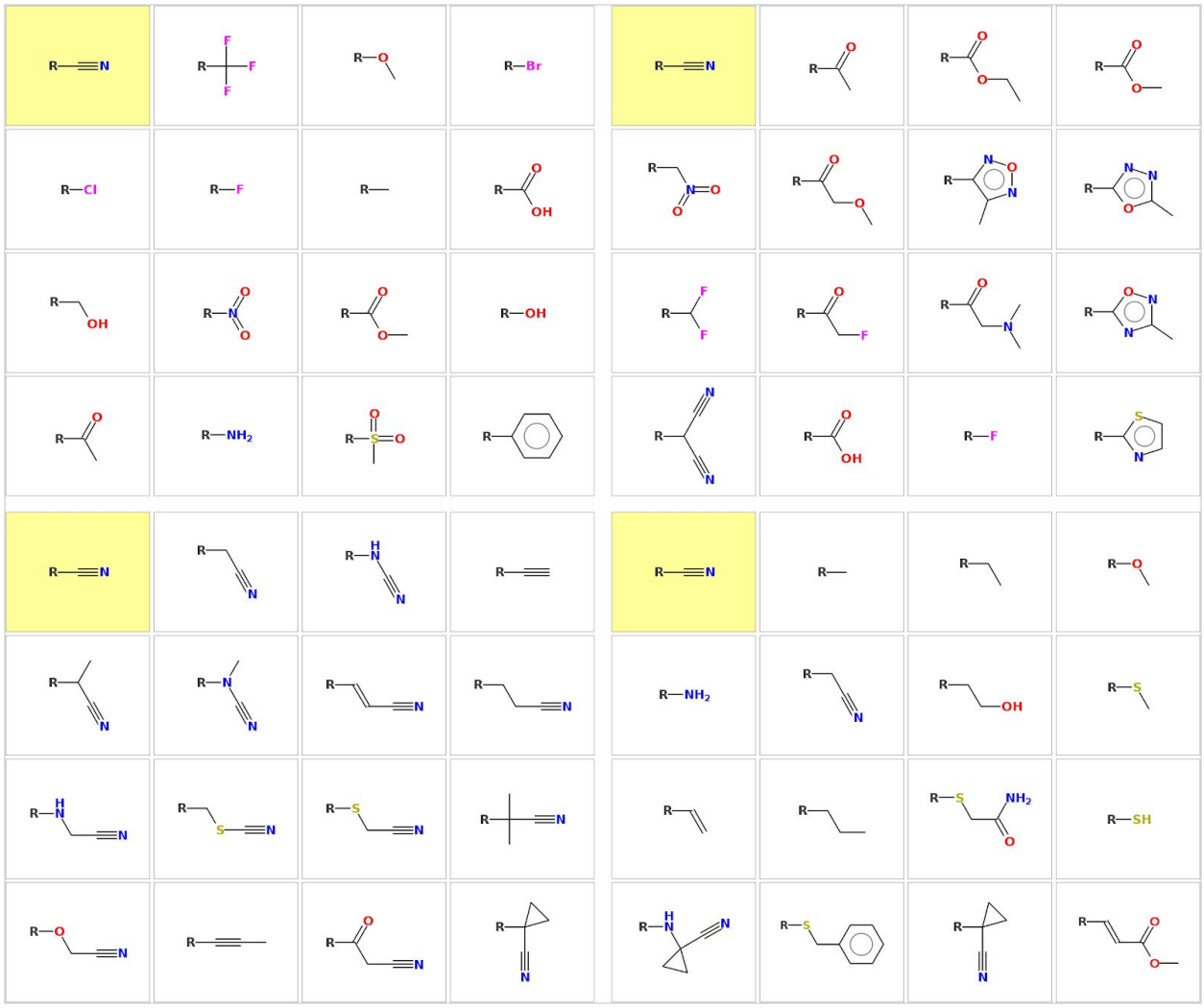
peter.ertl@novartis.com
www.peter-ertl.com

Snippet of the Python code showing the network architecture and the parameters used when training the model:

```
model = tf.keras.Sequential()
model.add(layers.Dense(10000,activation='relu'))
model.add(layers.Dense(10000,activation='relu'))
model.add(layers.Dense(ncy,activation='sigmoid'))

model.compile(optimizer=tf.keras.optimizers.Adam(0.01),loss='binary_crossentr
opy',metrics=['accuracy'])

model.fit(X,Y,epochs=30,batch_size=500,validation_split=0.2,verbose=2)

(X and Y are data matrices as described in the main text, ncy is the y
dimension of the Y matrix)
```

Comparison of 4 methods to identify bioisosteric analogs.

In the following figures the analogs of 4 simple substituents identified by 4 different search strategies are shown. Every figure consists of 4 parts corresponding to different search methods: top left - deep neural network (the method described in this article), top right - similarity in molecular physicochemical properties, bottom left - classical fingerprint similarity, bottom right - a null model described in details in the text of the article. One can compare these results also with the Figure 2 in the article where the bioisosteres of these 4 substituents extracted from medicinal chemistry literature are shown.


R—C≡N
R—CF3
R—O—CH3
R—Br
R—C≡N
R—C(=O)—CH3
R—C(=O)—O—CH2CH3
R—C(=O)—O—CH3
R—Cl
R—F
R—CH3
R—C(=O)—OH
R—CH2—N(=O)=O
R—C(=O)—O—CH3
R—C3H2O2N
R—C2H2O2N
R—CH2—OH
R—N(=O)=O
R—C(=O)—O—CH3
R—OH
R—CHF2
R—C(=O)—CH2F
R—C(=O)—N(CH3)2
R—C3H2O2N
R—C(=O)—CH3
R—NH2
R—S(=O)=O
R—C6H5
R—CH(C≡N)2
R—C(=O)—OH
R—F
R—C2H2NS
R—C≡N
R—CH2—C≡N
R—NH—C≡N
R—C≡CH
R—C≡N
R—CH3
R—CH2CH3
R—O—CH3
R—CH(CH3)—C≡N
R—N(CH3)—CH2—C≡N
R—CH=CH—C≡N
R—CH2—CH2—C≡N
R—NH2
R—CH2—C≡N
R—CH2—CH2—OH
R—S—CH3
R—NH—CH2—C≡N
R—CH2—S—CH2—C≡N
R—S—CH2—C≡N
R—C(CH3)(C≡N)—C≡N
R—CH=CH2
R—CH2—CH3
R—S—CH2—C(=O)—NH2
R—SH
R—O—CH2—C≡N
R—C≡C—CH3
R—C(=O)—CH2—C≡N
R—C(CH3)(C≡N)—C≡N
R—NH—C(CH2)—C≡N
R—S—C6H5
R—C(CH3)(C≡N)—C≡N
R—CH=CH—C(=O)—O—CH3