In [1]:
```
import os
import sys
import yaml
import pandas as pd
import numpy as np
import networkx as nx
import matplotlib.pyplot as plt
import glob
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import mean_squared_error
from math import sqrt
from arsenic import plotting, stats, wrangle, plotlying
%matplotlib inline
```

In [2]:
```
with open('example.csv','r') as f:
    data=f.readlines()
```

In [3]:
```
data
```

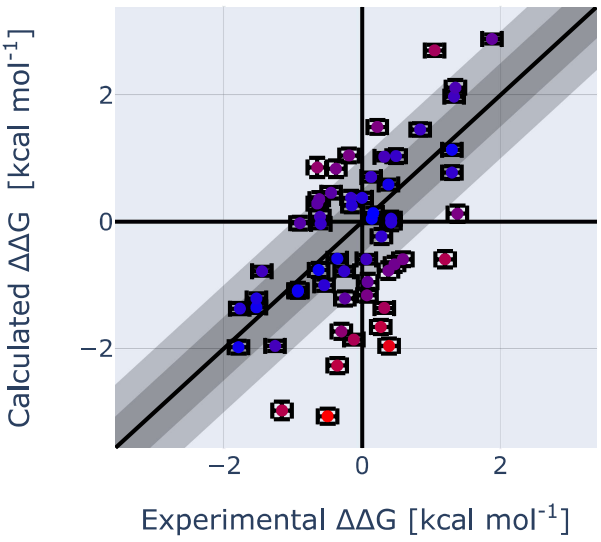
Out[3]:
```
['# Experimental block\n',
 '# Ligand, expt_DDG, expt_dDDG\n',
 'CAT-13a, -8.83 , 0.10 \n',
 'CAT-13b, -9.11 , 0.10\n',
 'CAT-13c, -9.31 , 0.10\n',
 'CAT-13d, -10.46, 0.10\n',
 'CAT-13e, -9.95 , 0.10\n',
 'CAT-13f, -9.08 , 0.10\n',
 'CAT-13g, -9.08 , 0.10\n',
 'CAT-13h, -9.62 , 0.10\n',
 'CAT-13i, -9.26 , 0.10\n',
 'CAT-13j, -8.72 , 0.10\n',
 'CAT-13k, -9.69 , 0.10\n',
 'CAT-13m, -8.75 , 0.10\n',
 'CAT-13n, -8.53 , 0.10\n',
 'CAT-13o, -8.53 , 0.10\n',
 'CAT-17a, -10.72, 0.10\n',
 'CAT-17b, -10.01, 0.10\n',
 'CAT-17c, -9.85 , 0.10\n',
 'CAT-17d, -9.41 , 0.10\n',
 'CAT-17e, -10.01, 0.10\n',
 'CAT-17f, -9.41 , 0.10\n',
 'CAT-17g, -9.73 , 0.10\n',
 'CAT-17h, -10.32, 0.10\n',
 'CAT-17i, -9.46 , 0.10\n',
 'CAT-24,  -11.34, 0.10\n',
 'CAT-4a,  -7.92 , 0.10\n',
 'CAT-4b,  -9.62 , 0.10\n',
 'CAT-4c,  -7.84 , 0.10\n',
 'CAT-4d,  -9.10 , 0.10\n',
 'CAT-4i,  -8.25 , 0.10\n',
 'CAT-4j,  -9.01 , 0.10\n',
 'CAT-4k,  -7.84 , 0.10\n',
 'CAT-4l,  -9.33 , 0.10\n',
 'CAT-4m,  -9.14 , 0.10\n',
 'CAT-4n,  -9.08 , 0.10\n',
 'CAT-4o,  -9.37 , 0.10\n',
 'CAT-4p,  -10.07, 0.10 \n',
 '\n',
 '# Calculated block\n',
 '# Ligand1,Ligand2, calc_DDG, calc_dDDG(MBAR), calc_dDDG(additional)\n',
 'CAT-13b,CAT-17g,\t0.36\t,0.11,0.0\n',
 'CAT-13a,CAT-17g,   -0.02\t,0.1,0.0\n',
 'CAT-13e,CAT-17g,    1.5\t,0.11,0.0\n',
 'CAT-4m ,CAT-4c ,   0.78\t,0.1,0.0\n',
 'CAT-13k,CAT-4d ,   -0.59\t,0.11,0.0\n',
 'CAT-24 ,CAT-17e,   1.98\t,0.08,0.0\n',
 'CAT-13g,CAT-17g,\t0.86\t,0.15,0.0\n',
 'CAT-13d,CAT-13h,   1.46\t,0.1,0.0\n',
 'CAT-13a,CAT-17i,\t-0.76\t,0.11,0.0\n',
 'CAT-4m ,CAT-13j,   -0.01\t,0.12,0.0\n',
 'CAT-13a,CAT-13m,   -0.95\t,0.13,0.0\n',
 'CAT-4l ,CAT-13k,\t-2.27\t,0.12,0.0\n',
 'CAT-13o,CAT-17i,\t-1.08\t,0.12,0.0\n',
 'CAT-4c ,CAT-4o ,\t-1.35\t,0.09,0.0\n',
 'CAT-4j ,CAT-4o ,\t-0.58\t,0.06,0.0\n',
 'CAT-4i ,CAT-13m,   -3.07\t,0.12,0.0\n',
 'CAT-24 ,CAT-17i,   2.89\t,0.07,0.0\n',
 'CAT-13j,CAT-4o ,\t0.29\t,0.1,0.0\n',
 'CAT-4n ,CAT-13k,\t0.08\t,0.1,0.0\n',
 'CAT-4o ,CAT-4b ,\t-1.21\t,0.11,0.0\n',
```

```
 'CAT-17i,CAT-13f,   -0.77\t,0.13,0.0\n',
 'CAT-17c,CAT-17e,\t0.38\t,0.11,0.0\n',
 'CAT-13k,CAT-4b ,   -1.16\t,0.1,0.0\n',
 'CAT-4m ,CAT-4j ,   0.71\t,0.07,0.0\n',
 'CAT-13n,CAT-13k,\t-2.98\t,0.14,0.0\n',
 'CAT-13d,CAT-17h,   0.05\t,0.07,0.0\n',
 'CAT-17b,CAT-13d,\t0.46\t,0.09,0.0\n',
 'CAT-4m ,CAT-4n ,   -0.59\t,0.07,0.0\n',
 'CAT-4m ,CAT-13k,\t-1.0    ,0.09,0.0\n',
 'CAT-13c,CAT-17i,\t0.26\t,0.11,0.0\n',
 'CAT-4a ,CAT-4o ,\t-0.78\t,0.07,0.0\n',
 'CAT-13h,CAT-17i,   0.14\t,0.09,0.0\n',
 'CAT-4o ,CAT-4d ,   -1.66\t,0.11,0.0\n',
 'CAT-17g,CAT-17c,\t-1.86\t,0.08,0.0\n',
 'CAT-4a ,CAT-13k,\t-1.37\t,0.09,0.0\n',
 'CAT-13d,CAT-17d,   2.71\t,0.09,0.0\n',
 'CAT-17g,CAT-17f,   -1.36\t,0.08,0.0\n',
 'CAT-13e,CAT-17i,   1.04\t,0.11,0.0\n',
 'CAT-13d,CAT-13b,   2.12\t,0.12,0.0\n',
 'CAT-17g,CAT-13i,   -0.67\t,0.12,0.0\n',
 'CAT-4m ,CAT-13m,   -1.96\t,0.12,0.0\n',
 'CAT-17g,CAT-13c,   0.05\t,0.11,0.0\n',
 'CAT-17i,CAT-17a,\t-1.96\t,0.08,0.0\n',
 'CAT-13d,CAT-13f,   0.13\t,0.13,0.0\n',
 'CAT-17f,CAT-17e,   -0.03\t,0.07,0.0\n',
 'CAT-13d,CAT-17a,\t-0.78\t,0.07,0.0\n',
 'CAT-17g,CAT-17d,   1.03\t,0.06,0.0\n',
 'CAT-13n,CAT-13a,   -1.73\t,0.12,0.0\n',
 'CAT-13o,CAT-17h,\t-1.98\t,0.1,0.0\n',
 'CAT-17b,CAT-17e, 0.38\t,0.08,0.0\n',
 'CAT-4k ,CAT-4o ,\t-1.21\t,0.08,0.0\n',
 'CAT-4m ,CAT-4l ,\t1.05\t,0.11,0.0\n',
 'CAT-4m ,CAT-4k ,   1.14   , 0.08,0.0\n',
 'CAT-13n,CAT-4i ,   -0.23\t,0.12,0.0\n',
 'CAT-13g,CAT-17i,\t0.84\t,0.13,0.0\n',
 'CAT-4p ,CAT-13k,   0.59\t,0.08,0.0\n',
 'CAT-4m ,CAT-4p ,\t-1.1\t,0.06,0.0\n',
 'CAT-13d,CAT-13i,  -0.59   , 0.13,0.0\n']
```

In [4]:
```
fe = wrangle.FEMap('example.csv')
```

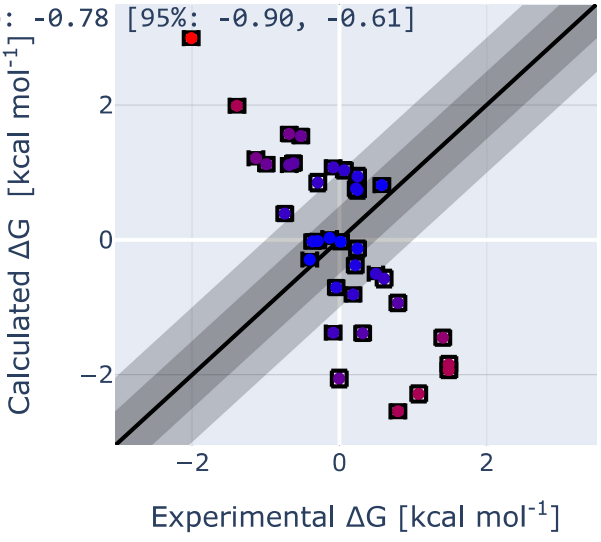
In [5]:
```
plotting.plot_DDGs(fe.graph,target_name='target', title='method',plotly=True)
```

```
method
target (N = 58)
RMSE: 1.05 [95%:  0.88,  1.21]
MUE:  0.87 [95%:  0.71,  1.02]
```

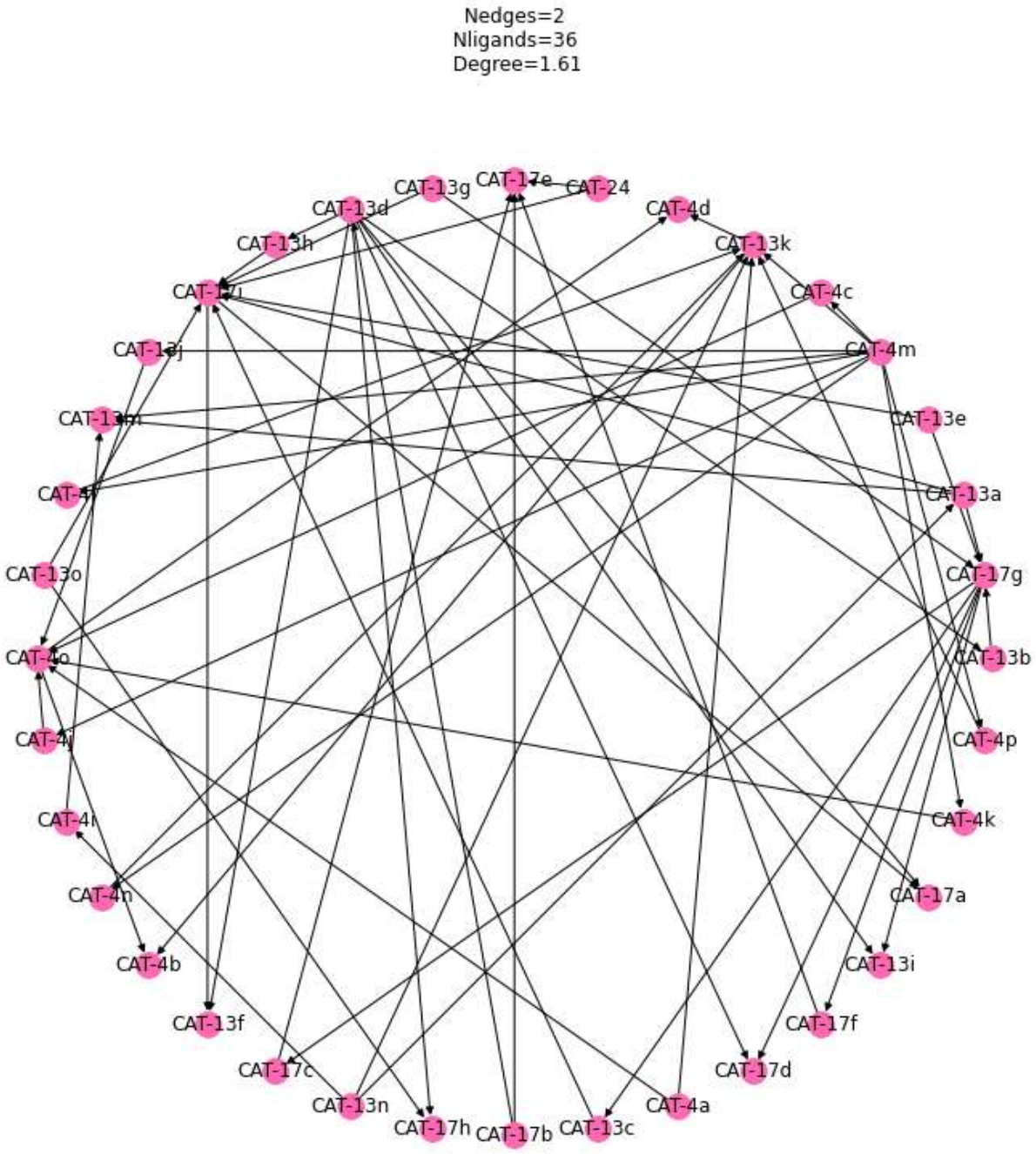


In [6]:
```
plotting.plot_DGs(fe.graph,method_name='',target_name='target',title='method',guidelines
=True,plotly=True)
```

```
method
target (N = 36)
RMSE: 1.96 [95%:  1.48,  2.37]
MUE:  1.55 [95%:  1.18,  1.93]
R2:   0.61 [95%:  0.38,  0.79]
rho: -0.78 [95%: -0.90, -0.61]
```



In [7]:
```
fe.draw_graph()
```



In [8]:
```
for node in fe.graph.nodes(data=True):
    print(node[1])
```

{'name': 'CAT-13b', 'exp_DG': -9.11, 'exp_dDG': 0.1, 'calc_DG': -0.378864096343817, 'ca
lc_dDG': 0.10298265270254682}
{'name': 'CAT-17g', 'exp_DG': -9.73, 'exp_dDG': 0.1, 'calc_DG': -0.2937193764396495, 'c
alc_dDG': 0.06974040707254545}
{'name': 'CAT-13a', 'exp_DG': -8.83, 'exp_dDG': 0.1, 'calc_DG': -0.5032132305330059, 'c
alc_dDG': 0.06634945975707236}
{'name': 'CAT-13e', 'exp_DG': -9.95, 'exp_dDG': 0.1, 'calc_DG': 1.1368899147871745, 'ca
lc_dDG': 0.09905155826710046}
{'name': 'CAT-4m', 'exp_DG': -9.14, 'exp_dDG': 0.1, 'calc_DG': -0.8106628346268927, 'ca
lc_dDG': 0.08959799229915309}
{'name': 'CAT-4c', 'exp_DG': -7.84, 'exp_dDG': 0.1, 'calc_DG': -1.848626634225056, 'cal
c_dDG': 0.11092290306508958}
{'name': 'CAT-13k', 'exp_DG': -9.69, 'exp_dDG': 0.1, 'calc_DG': -0.023929983396378418,
'calc_dDG': 0.09107943588485143}
{'name': 'CAT-4d', 'exp_DG': -9.1, 'exp_dDG': 0.1, 'calc_DG': 0.7592463523520268, 'calc
_dDG': 0.11704611431469873}
{'name': 'CAT-24', 'exp_DG': -11.34, 'exp_dDG': 0.1, 'calc_DG': 2.9925206635254504, 'ca
lc_dDG': 0.08542219157291443}
{'name': 'CAT-17e', 'exp_DG': -10.01, 'exp_dDG': 0.1, 'calc_DG': 1.1105078733362972, 'c
alc_dDG': 0.08270317963084474}
{'name': 'CAT-13g', 'exp_DG': -9.08, 'exp_dDG': 0.1, 'calc_DG': 0.7382963115097869, 'ca
lc_dDG': 0.11506211165100412}
{'name': 'CAT-13d', 'exp_DG': -10.46, 'exp_dDG': 0.1, 'calc_DG': 1.2113768981503923, 'c
alc_dDG': 0.0754493503773348}
{'name': 'CAT-13h', 'exp_DG': -9.62, 'exp_dDG': 0.1, 'calc_DG': -0.01872127430064019,
'calc_dDG': 0.09368187647268233}
{'name': 'CAT-17i', 'exp_DG': -9.46, 'exp_dDG': 0.1, 'calc_DG': 0.027499206014022448,
'calc_dDG': 0.0707608398760847}
{'name': 'CAT-13j', 'exp_DG': -8.72, 'exp_dDG': 0.1, 'calc_DG': -0.5745795753124074, 'c
alc_dDG': 0.1170882391245374}
{'name': 'CAT-13m', 'exp_DG': -8.75, 'exp_dDG': 0.1, 'calc_DG': 0.8112617788559842, 'ca
lc_dDG': 0.08852246862194565}
{'name': 'CAT-4l', 'exp_DG': -9.33, 'exp_dDG': 0.1, 'calc_DG': -2.0584942497254124, 'ca
lc_dDG': 0.11752762883991245}
{'name': 'CAT-13o', 'exp_DG': -8.53, 'exp_dDG': 0.1, 'calc_DG': -0.9375574913660545, 'c
alc_dDG': 0.10655453679098983}
{'name': 'CAT-4o', 'exp_DG': -9.37, 'exp_dDG': 0.1, 'calc_DG': -0.707577311899569, 'cal
c_dDG': 0.0960579000082212}
{'name': 'CAT-4j', 'exp_DG': -9.01, 'exp_dDG': 0.1, 'calc_DG': -1.3862958862311414, 'ca
lc_dDG': 0.10036834627044465}
{'name': 'CAT-4i', 'exp_DG': -8.25, 'exp_dDG': 0.1, 'calc_DG': -2.286255019779384, 'cal
c_dDG': 0.10974651000139549}
{'name': 'CAT-4n', 'exp_DG': -9.08, 'exp_dDG': 0.1, 'calc_DG': -0.12965672918866833, 'c
alc_dDG': 0.10361013218319477}
{'name': 'CAT-4b', 'exp_DG': -9.62, 'exp_dDG': 0.1, 'calc_DG': 0.8493517684121379, 'cal
c_dDG': 0.1145407401065681}
{'name': 'CAT-13f', 'exp_DG': -9.08, 'exp_dDG': 0.1, 'calc_DG': 0.9394380520821999, 'ca
lc_dDG': 0.11206157396793683}
{'name': 'CAT-17c', 'exp_DG': -9.85, 'exp_dDG': 0.1, 'calc_DG': 1.5400673478071627, 'ca
lc_dDG': 0.09264651595657337}
{'name': 'CAT-13n', 'exp_DG': -8.53, 'exp_dDG': 0.1, 'calc_DG': -2.5437718184147466, 'c
alc_dDG': 0.0892160917320019}
{'name': 'CAT-17h', 'exp_DG': -10.32, 'exp_dDG': 0.1, 'calc_DG': 1.1222642465644506, 'c
alc_dDG': 0.09416806960653072}
{'name': 'CAT-17b', 'exp_DG': -10.01, 'exp_dDG': 0.1, 'calc_DG': 1.5703397187714856, 'c
alc_dDG': 0.09232047394355672}
{'name': 'CAT-13c', 'exp_DG': -9.31, 'exp_dDG': 0.1, 'calc_DG': -0.02811008521280911,
'calc_dDG': 0.09905155826710019}
{'name': 'CAT-4a', 'exp_DG': -7.92, 'exp_dDG': 0.1, 'calc_DG': -1.4522794726945252, 'ca
lc_dDG': 0.10502862955757894}
{'name': 'CAT-17d', 'exp_DG': -9.41, 'exp_dDG': 0.1, 'calc_DG': -1.3775359073350213, 'c

```
alc_dDG': 0.08213276017128228}
{'name': 'CAT-17f', 'exp_DG': -9.41, 'exp_dDG': 0.1, 'calc_DG': 1.0743385349378758, 'ca
lc_dDG': 0.08722526835297208}
{'name': 'CAT-13i', 'exp_DG': -9.26, 'exp_dDG': 0.1, 'calc_DG': 1.0319159703365965, 'ca
lc_dDG': 0.1083541906599787}
{'name': 'CAT-17a', 'exp_DG': -10.72, 'exp_dDG': 0.1, 'calc_DG': 1.989695421029305, 'ca
lc_dDG': 0.08519034807920653}
{'name': 'CAT-4k', 'exp_DG': -7.84, 'exp_dDG': 0.1, 'calc_DG': -1.9341200732632249, 'ca
lc_dDG': 0.1050229267217822}
{'name': 'CAT-4p', 'exp_DG': -10.07, 'exp_dDG': 0.1, 'calc_DG': 0.388960991816095, 'cal
c_dDG': 0.09895523959628864}
```

In [ ]: