A Computational Model for Understanding Stem Cell, Trophectoderm and Endoderm Lineage Determination

Introduction

Embryonic stem cells can differentiate into distinct cell type. The aim is to understand how to reprogram these differentiated cells back to the embryonic stem cell, to do so one must understand the mechanism behind it. There is a subset of transcription factors (TF) that are responsible for (1) the initial differentiation and/or maintenance of a steady state and subsequently (2) the possible reprogram back to self-renewing and pluripotent stem cells. A mathematical model on lineage determination is possible to derive due to the limited number of TF involved in the process and the functional quantification of the system via ChIP-on-chip and microarray analysis. The model could simulate new scenarios, such as reprograming, and aid us with understanding the behaviours expressed by the system.

Lineage determination (e.g. trophectoderm and endoderm) and conservation of the stem cell state is determined by the stable state the system tends to. The result of the system's steady state is determined by the dynamics and dependencies of the TF and their concentrations (Fig.1). Oct4, Sox2 and the Nanog TF are responsible for the maintenance of the embryonic stem cells state. It's suggested that the expression of these factors maintain the pluripotency and self-renewal character in cells, while their change in expression (higher or lower) causes cells to differentiate. This dynamic proposes that these three factors function as a bistable switch.

A high expression of Cdx2 and a low expression of Oct4 leads the trophectoderm lineage. Conversely, a high expression of Oct4 keeps the stem cell stage. Nanog and Gata-6 relationship determines if the embryonic stem cell enters the endoderm lineage. A high expression of Gata-6 differentiates the cell, whereas Nanog's expression keeps the stem cell state.

In microarray experiments that evaluated the expression of genes at varying concentration of Oct4 show bell/inversed bell-shaped concentration curves. The cells in the experiment expressed differentiating factors at low and high concentrations of Oct4. Differentiation did not seem to occur at mid-range concentrations of Oct4 therefore cells remained in the embryonic stem cell state. The stable stem cell state is maintained by the mid-range concentration of Oct4, Nanog and Sox2. Low concentration of Oct4 lead cells into the trophectoderm lineage, whereas high expression leads to the endoderm lineage. Any significant deviation in Oct4, Nanog and Sox2 concentration will cause the cell to differentiate.

The architecture and dynamics of the system causes it to have three steady states: endoderm, trophectoderm and embryonic stem cell. To be able to move among these steady states, specific TFs need to be either expressed or under expressed to change the type of cell.

In this paper I will present and discuss the mutually antagonistic expression of two TF, Nanog and Gata-6 as a function of the concentration of Oct4. We will also observe the progress of the system when other TF (Cdx2, Gata-6 and Gcnf) are accounted for. Ultimately, based on the observations of the two initial graphs I will investigate the possibility of returning to the stem cell state with a different composition of transcription factor concentration. I expect

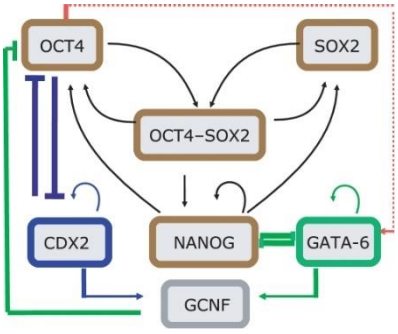


Figure 1- Transcription Factor Interactions which determines the system's lineage.

that this alternate bistable state will conserve the pluripotent characteristic of stem cells.

Results

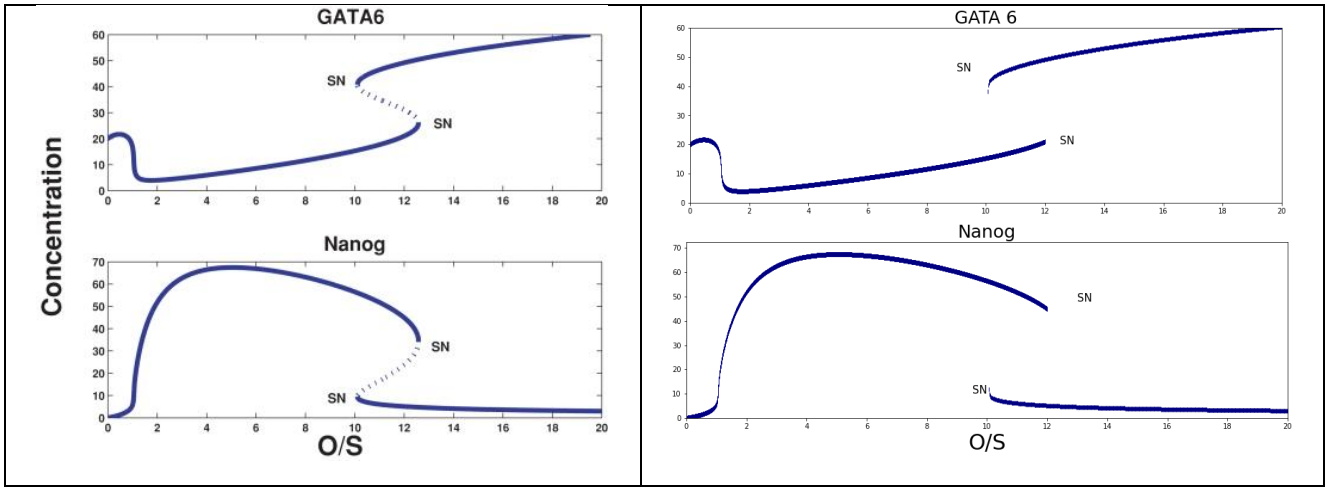


Figure 2. ***Left: Published model. Right: Reproduced model****. Steady state values of Nanog and Gata-6 as functions of the concentration of Oct4-Sox4 (O/S). GATA6 vs O/S depicts an inverted bell-curve while Nanog vs O/S is a bell-curve. Two saddle nodes (SN) are present in both graphs indicating the presence of a bistable state or hysteresis. The graph depicts the system's repression character between Nanog and Gata6, as their concentration is inversely proportional. The space between the SN is the unstable state.*

The interaction between the regulatory factors Nanog, Gata-6 and Oct4 determine the fate of the cell via a network like mechanism with biphasic behaviour. O/S represents the hetero-dimer Oct4-Sox4. O/S is shown to activate both Nanog and Gata-6 but at different concentrations (Fig.2). Nanog's high expression represses the expression of Gata-6. However, at low initial concentrations of Nanog, Nanog's Gata-6-repressing-effect is turned off and since Gata-6 is self-activating it will experience a slight high. At low levels of O/S, Nanog is expressed. After a certain threshold, O/S begins to form a heterodimer with Gata-6 which inhibits the production of Nanog and thus the repression of Gata-6.

The system depicts hysteretic behaviour. The cell is presumably in a steady stem cell state at low concentrations of O/S where Nanog is highly expressed. Once the O/S concentration passes a threshold (saddle node), the system finds a new steady state at the endoderm lineage (with high Gata-6). To revert to the stem cell state, the system's O/S needs to drop to a second saddle node, which induces Nanog production, prevents Gata-6's self-activation and the formation of the Nanog repressing O/S-Gata-6 heterodimer.

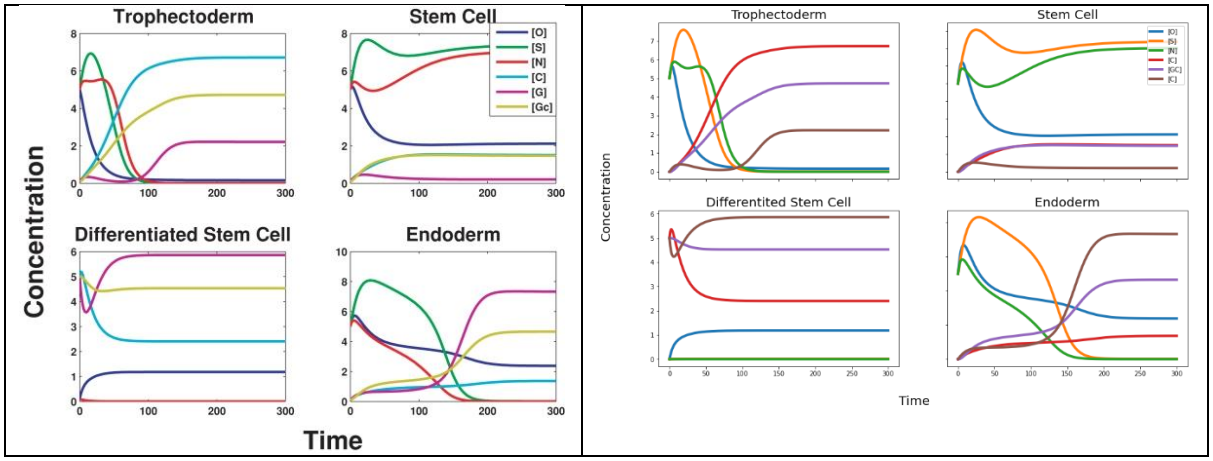


Figure 3 - ***Shows the concentration levels of 6 stem cell transcription factors over time tending towards 3 different steady states.*** *The trophectoderm, stem cell and endoderm lineage are products of a high, moderate, and low quantity of factor A, respectively. These three states commence with high initial values of Oct4 [O], Sox2 [S] and Nanog [N] compared to Cdx2 [C], Gata-6 [G] and Gcnf [Gc] which are initially absent. The differentiated*

stem cell lineage has low [O], [S], [N], and high [C], [G], [Gc] initial values. The system is by stable at moderate levels of A as two states exist differentiated and non-differentiated stem cells. The bistability is determined by the initial concentration of TF.

The system is based on factor A, it is an arbitrary factor whose concentration controls the production of Oct4, and subsequently the steady state the system holds.

The trophectoderm state is achieved at low levels of A (A=1). Oct4 is inactivated and consequently so are Sox2 and Nanog. This leads to the increase of Cdx2 and Gata-6 due to their self-activating nature. Cdx2 then stimulates Gcnf activation, which is a known inhibitor of Oct4. Similarly, Gata-6 inhibits the activation of Nanog. This causes the system to achieve the trophectoderm steady state from an original stem cell state with a high Oct4, Sox2 and Nanog concentration.

The endoderm state is achieved at high levels of A (A=25). Oct4 is highly activated at high levels of A, accordingly, as depicted in Figure 1, high levels of Oct4 activates Gata-6 production, which is self-activating and a Nanog repressor. The repression of Nanog prevents the minimal positive activation of Oct4 and Sox2, causing a subtle decrease in Oct4 and complete inactivation of Sox2. Cdx2 is self-activating and simultaneously repressed by Oct2. The moderate amount of Cdx2 and high Gata-6 then stimulates Gcnf activation leading to its increase. Oct4, Gcnf and Gata-6 activation causes the endoderm steady state in the system from high Oct4, Sox2 and Nanog concentrations.

The system at moderate A concentration (A=10) is bistable and is dependent on the difference of the initial concentration of transcription factors. Cells that commence with higher concentrations of Oct4, Sox2 and Nanog compared to Gcnt, Gata-6 and Cdx2 keep the system in the stem cell state, it can be described as the ‘on’ state of the switch. However, when the concentrations are reversed, the system experiences the differentiated stem cell state, the “other-on” state. This shows the bistable nature at a moderate expression of Oct4 by a moderate inducing factor (A).

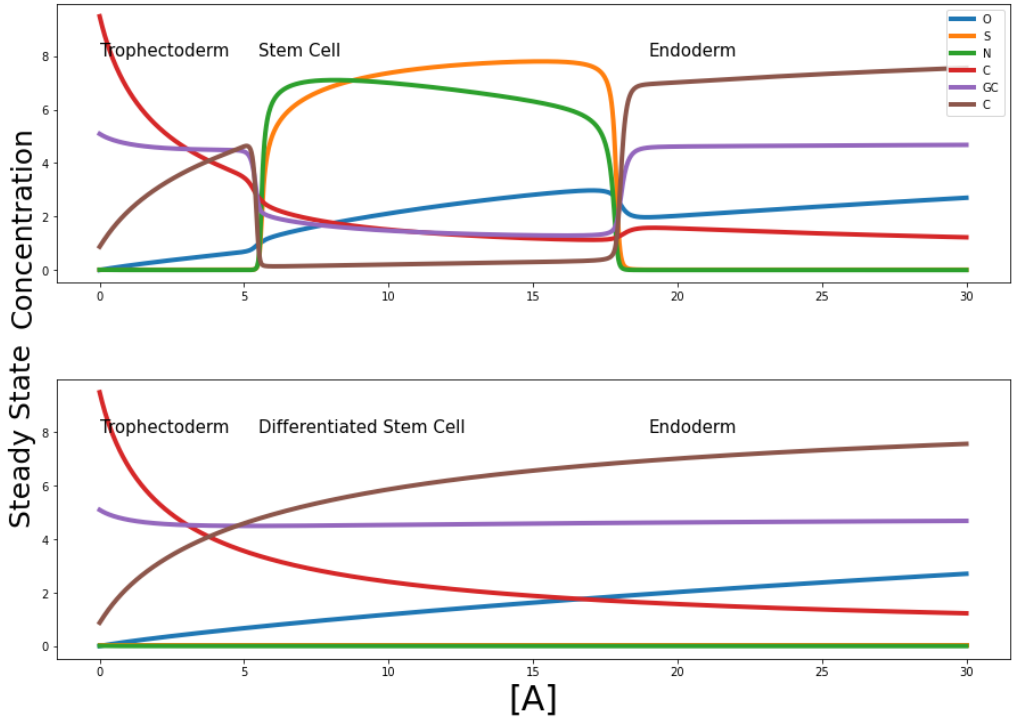


*Figure 4 - **The steady state concentration of the transcription factors at changing concentrations of A.** The top figure shows the fate of the system when Oct4, Sox2 and Nanog are higher in concentration compared to*

Gcnf, Cdx2 and Gata-6. The bottom figure shows the fate of the system when Gcnf, Cdx2 and Gata-6 are higher in concentration compared to Oct4, Sox2 and Nanog

The figures depict a collective model of what would occur in a system/cell if a single factor (A) was progressively changed. All cells commence as stem cells with high concentration of Oct4, Sox2 and Nanog. The change in these transcription factors, primarily regulated by Oct4, would cause the system to remain in the stem cell state ("on") or become endodermal or trophectodermal ("off"). If the system begins as a stem cell with higher concentrations of Oct4, Sox2 and Nanog (Fig.4 top) it imitates how cells naturally differentiate.

A differentiated cell commencing with a high concentration of Gcnf, Cdx2 and Gata-6 (Fig. 4 bottom), would return to a steady state that has a low concentration of Oct4, Sox2 and Nanog. Based on the model, Figure 4 suggests that from this "other" differentiated stem cell state, the system is able to, once again, differentiate into the trophectoderm or endoderm state. This suggest that the system does not have to return to the original stem cell state, with high Oct4, Sox2 and Nanog, to exhibit behaviours of pluripotency or self-renewal. Because of the bistability of the system these behaviours can be achieved at other concentrations of the TF as suggested by the mathematical model.

Discussion

The system that coordinates the fate of stem cells is govern by the interactions and concentration of a few TF. A model was produced to mimic the dynamics for this system as observed in microarray and ChiP-on-chip experiments. The primary TF, Oct4, is responsible for many of the interactions occurring in the system. Its concentration seems to be the determining factor of the fate of the cell as seen in its switch like inducing behaviour while interacting with Nanog and gata-6. This would suggest that Oct4 will play an important role in future experiments regarding the return to stem-cell like status.

The model was able to express the concentrations of different TF as the system tended towards differentiation. The trophectoderm state was achieved when Oct4, Nanog and Sox2 were inactivated and Cdx2, Gata-6 and Gcnf were activated. However, the over expression of Oct4 lead the system to pursue the endoderm state. If the system was kept at elevated levels of Oct4, Nanog and Sox2, it would remain in the stem cell state.

Curiously the model suggested that the two stem-cell states could potentially give the system a pluripotent characteristic. Perhaps the cell does not have to be retuned to a stem-cell that has elevated levels of Oct4, Nanog and Sox2, but with Cdx2, Gata-6 and Gcnf. This is an ambitious assumption; however the model does depict it. Further experiments altering the TF concentration should explore the possibility of achieving the stem cell state with the aid of Cdx2, Gata-6 and Gcnf.

Chickarmane, V., & Peterson, C. (2008). A computational model for understanding stem cell, trophectoderm and endoderm lineage determination. *PloS one*, *3*(10), e3478. https://doi.org/10.1371/journal.pone.0003478

In [1]:
```
import numpy as np
from scipy.integrate import odeint
%matplotlib inline
import matplotlib.pyplot as plt
```

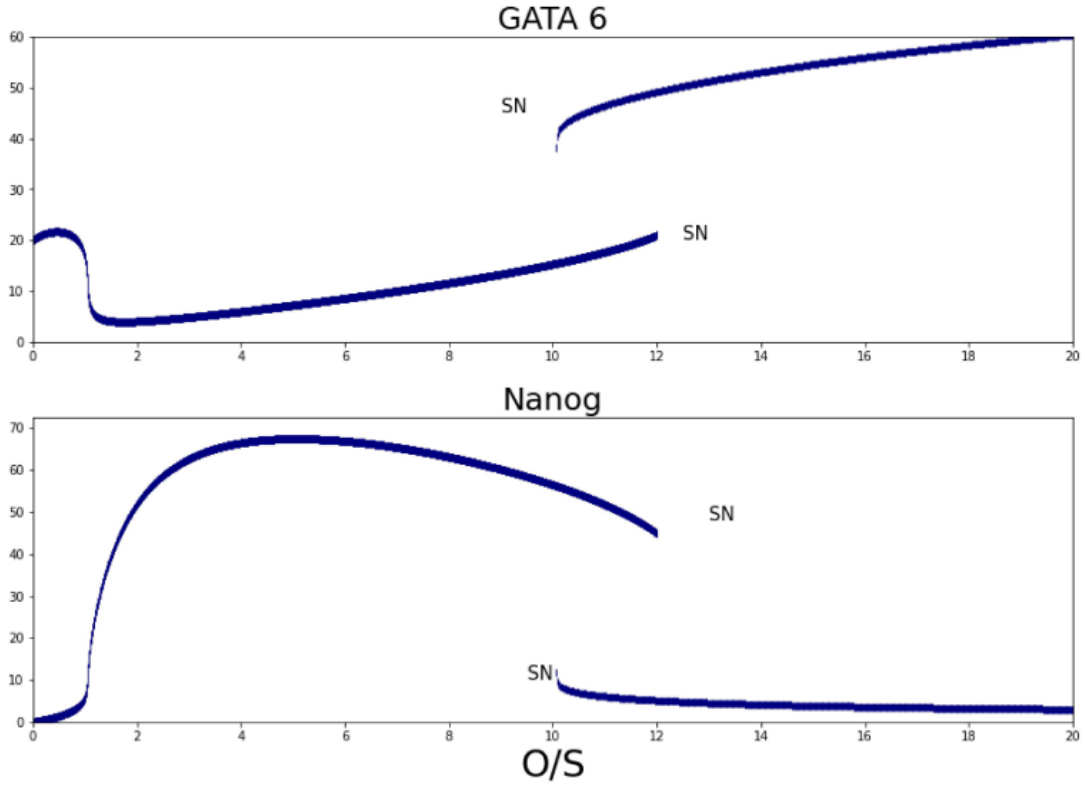
In [2]:
```
def NG (y, t, OS ):
    a1 = 0.02
    a2 = 0.0125
    b1 = 0.02
    b2 = 0.0125
    b3 = 0.03
    yn = 0.01
    c1 = 0.05
    c2 = 0.0125
    d1 = 0.05
    d2 = 0.0125
    d3 = 0.05
    yg = 0.01
    dydt = np.empty(2) #0 Nanog, 1 Gata
    dydt[0] = ((a1*OS+a2*OS*y[0])/(1+b1*OS+b2*OS*y[0]+b3*OS*y[1]))-yn*y[0]
    dydt[1] = ((c1*OS+c2*y[1])/(1+d1*OS+d2*y[1]+d3*y[0]))-yg*y[1]
    return(dydt)
```

In [3]:
```
t = np.linspace(0,10000)
OS= np.concatenate(((np.linspace(0,12,2000)),(np.linspace(30,10,1000))))
```

In [4]:
```
ansN = list()
yi = np.array([0,20])
for i in range(len(OS)):
    ans = odeint(NG,yi, t, args=(OS[i],))
    ansN.append(ans[-1])
    yi = ans[-1]

N = np.zeros(len(ansN))
G= np.zeros(len(ansN))

for i in range(len(ansN)):

    N[i] = ansN[i][0]
    G[i] = ansN[i][1]
```

In [5]:
```
plt.figure(figsize=(15,4.5))
plt.xticks([0,2,4,6,8,10,12,14,16,18,20])
plt.plot(OS,G,"|", color = "navy")
plt.title("GATA 6", fontsize=25)
plt.text(12.5, 20, 'SN',fontsize=15)
plt.text(9, 45, 'SN',fontsize=15)
plt.xlim(0,xmax=20)
plt.ylim(0, 60)
plt.show()

plt.figure(figsize=(15,4.5))
plt.xticks([0,2,4,6,8,10,12,14,16,18,20])
plt.plot(OS,N,"|",  color = "navy")
plt.title("Nanog", fontsize=25)
plt.xlabel("O/S", fontsize=30)
```

```
plt.text(9.5, 10, 'SN',fontsize=15)
plt.text(13, 48, 'SN',fontsize=15)
plt.xlim(0,xmax=20)
plt.ylim(0, max(N)+5)
plt.show()
```



In [6]:

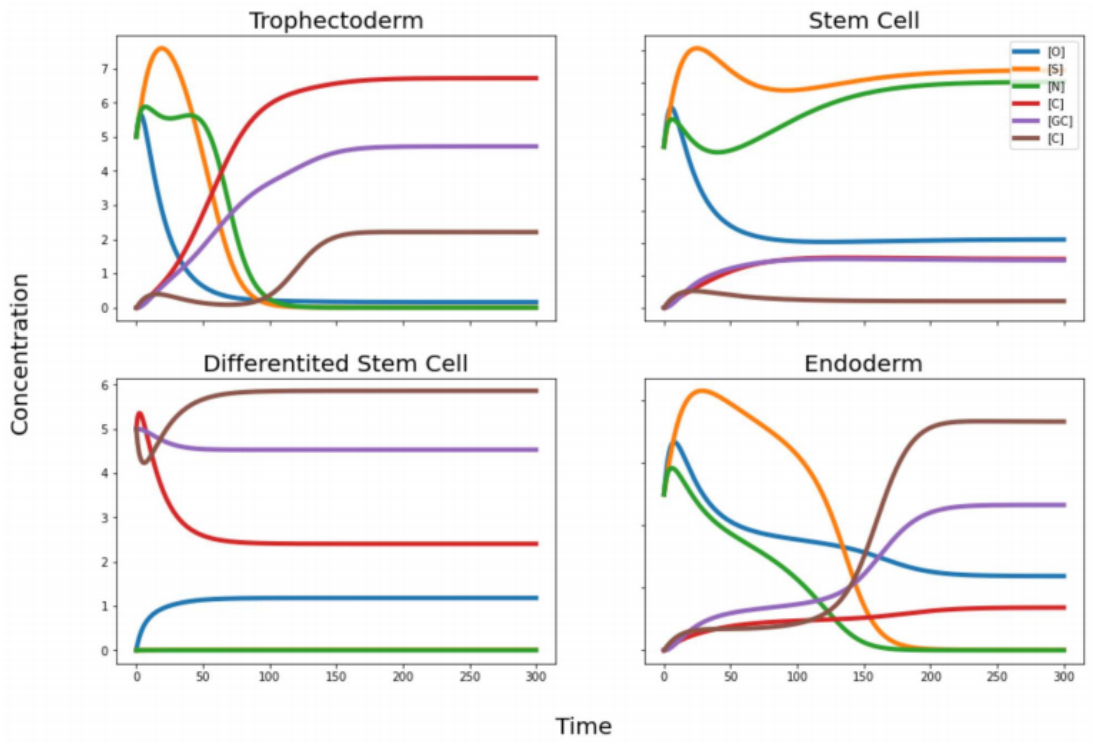
```
def Sys_ (y, t, A):
    a0, a1, a2, a3 = 0.001, 1, 0.005, 0.025
    b0, b1, b2, b3, b4, b5 = 1, 0.001, 0.005, 0.025,10, 10
    c0, c1, c2 = 0.001, 0.005, 0.025
    d0, d1, d2 =0.001, 0.005, 0.025,
    e0, e1, e2 = 0.001,0.1,0.1
    f0, f1, f2, f3 = 0.001,0.1,0.1,10
    g0, g1 = 0.001, 2
    h0, h1 = 2,5
    i0, i1, i2 = 0.001,0.1,0.1
    j0, j1 = 0.1,0.1
    p0, p1, p2 = 0.1,1,0.00025
    q0, q1, q2 = 1,0.00025, 15
    y1 = y2= y3= y4= y5= yg = 0.1

    dydt = np.empty(6) #0 O, #1 S, #2 N, #3 C, #4 GC, #5 G
    O,S,N,C,GC,G= y[0],y[1],y[2],y[3],y[4],y[5]

    dydt[0] = (a0 + a1*A + a2*O*S + a3*O*S*N)/(1 + b0*A + b1*O + b2*O*S + b3*O*S*N + b4
    dydt[1] = (c0 + c1*O*S + c2*O*S*N)/(1 + d0*O + d1*O*S + d2*O*S*N) - y2*S
    dydt[2] = (e0 + e1*O*S + e2*O*S*N)/(1 + f0*O + f1*O*S + f2*O*S*N + f3*O*G) - y3*N
    dydt[3] = (g0 + g1*C)/(1 + h0*C + h1*C*O) - y4*C
    dydt[4] = (i0 + i1*C + i2*G)/(1 + j0*C + j1*G) - y5*GC
    dydt[5] = (p0 + p1*O + p2*G)/(1 + q0*O + q1*G + q2*N) - yg*G
```

```
    return(dydt)
```

In [7]:
```
t = np.arange(301)
A= np.array([1,10,10,25])
```

In [8]:
```
Gr_ans = list()

for i in range(len(A)):
    O = list()
    S = list()
    N = list()
    C = list()
    GC = list()
    G = list()
    yi = np.array([5,5,5,0,0,0])
    if i == 2:
        yi = np.array([0,0,0,5,5,5])
    ans = odeint(Sys_,yi, t, args=(A[i],))
    Gr_ans.append(ans)
```

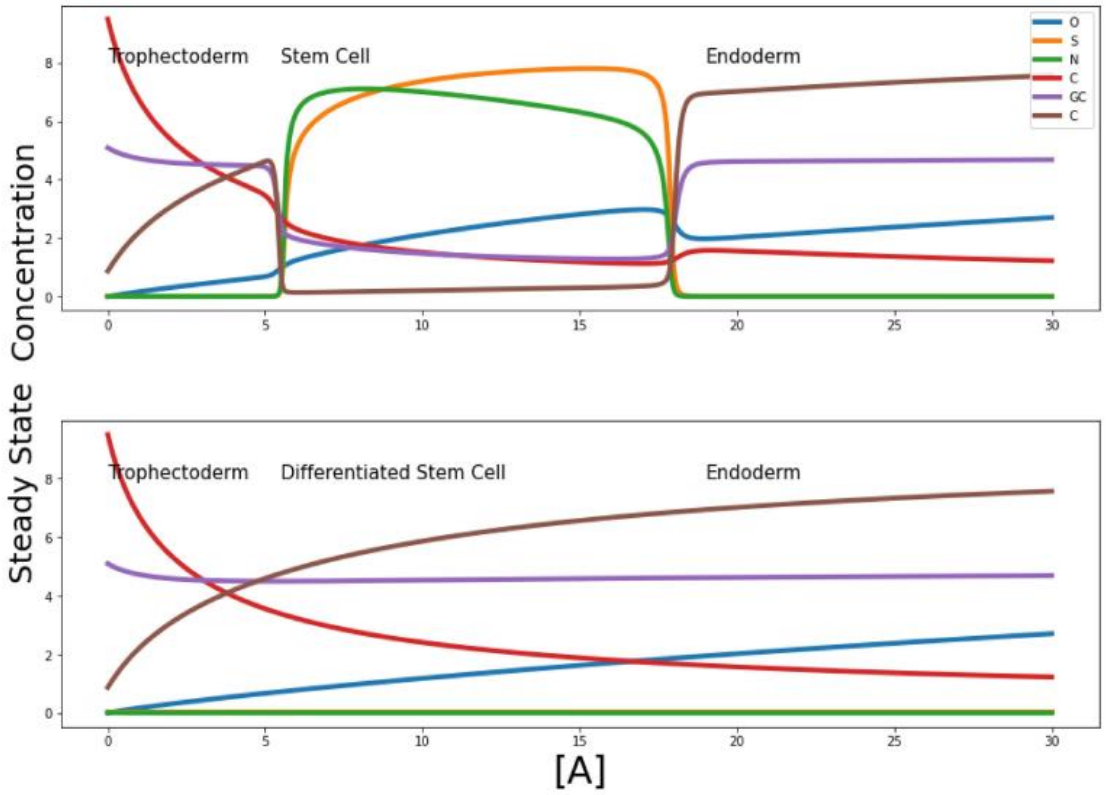
In [9]:
```
fig, axs = plt.subplots(2, 2,figsize=(15,10))
axs[0, 0].plot(t, Gr_ans[0],linewidth=4.0)
axs[0, 0].set_title('Trophectoderm',fontsize=20)
axs[0, 1].plot(t, Gr_ans[1],linewidth=4.0)
axs[0, 1].set_title('Stem Cell',fontsize=20)
axs[0, 1].legend(['[O]','[S]','[N]','[C]', '[GC]','[C]'],loc ="upper right")
axs[1, 0].plot(t, Gr_ans[2],linewidth=4.0)
axs[1, 0].set_title('Differentited Stem Cell',fontsize=20)
axs[1, 1].plot(t, Gr_ans[3],linewidth=4.0)
axs[1, 1].set_title('Endoderm',fontsize=20)


fig.text(0.5, 0.04, 'Time', ha='center',fontsize=20)
fig.text(0.04, 0.5, 'Concentration', va='center', rotation='vertical',fontsize=20)

for ax in axs.flat:
    ax.label_outer()
```



In [10]:
```
A= np.linspace(0,30,1000)
B = A[::-1]
t = np.arange(301)
```

In [11]:
```
Gr_ans = list()
Gr_ans2 = list()
for i in range(len(A)):
    yi = np.array([5,5,5,0,0,0])
    ans = odeint(Sys_,yi, t, args=(A[i],))
    Gr_ans.append(ans[-1])
    yj = np.array([0,0,0,5,5,5])
    ans = odeint(Sys_,yj, t, args=(A[i],))
    Gr_ans2.append(ans[-1])
```

In [12]:
```
plt.figure(figsize=(15,4.5))
plt.plot(A,Gr_ans,linewidth=4.0)
plt.text(0, 8, 'Trophectoderm',fontsize=15)
plt.text(5.5, 8, 'Stem Cell',fontsize=15)
plt.text(19, 8, 'Endoderm',fontsize=15)
plt.ylabel("Concentration", fontsize=25,position=(0.5,0.2))

plt.legend(['O','S','N','C', 'GC','C'],loc ="upper right")
plt.show()

plt.figure(figsize=(15,4.5))
plt.plot(A,Gr_ans2,linewidth=4.0)
plt.text(0, 8, 'Trophectoderm',fontsize=15)
plt.text(5.5, 8, 'Differentiated Stem Cell',fontsize=15)
plt.text(19, 8, 'Endoderm',fontsize=15)
plt.xlabel("[A]", fontsize=30)
plt.ylabel("Steady State", fontsize=25,position=(0.5,0.8))
plt.show()
```


Steady State Concentration
Trophectoderm
Stem Cell
Endoderm
O
S
N
C
GC
C
Trophectoderm
Differentiated Stem Cell
Endoderm
[A]