# CONTENTS

Preface v
A Note to the Reader vii
List of Reviewers xii
Contents xiii

## Part 1 Background Basics

### Chapter 1 The Nucleic Acid World

#### 1.1 The Structure of DNA and RNA
- DNA is a linear polymer of only four different bases 5
- Two complementary DNA strands interact by base pairing to form a double helix 7
- RNA molecules are mostly single stranded but can also have base-pair structures 9

#### 1.2 DNA, RNA, and Protein: The Central Dogma
- DNA is the information store, but RNA is the messenger 10
- Messenger RNA is translated into protein according to the genetic code 12
- Translation involves transfer RNAs and RNA-containing ribosomes 13

#### 1.3 Gene Structure and Control
- RNA polymerase binds to specific sequences that position it and identify where to begin transcription 14
- The signals initiating transcription in eukaryotes are generally more complex than in bacteria 15
- Eukaryotic mRNA transcripts undergo several modifications prior to their use in translation 17
- The control of translation 18

#### 1.4 The Tree of Life and Evolution
- A brief survey of the basic characteristics of the major forms of life 20
- Nucleic acid sequences can change as a result of mutation 21

Summary 23
Further Reading 24

### Chapter 2 Protein Structure

#### 2.1 Primary and Secondary Structure
- Protein structure can be considered on several different levels 25
- Amino acids are the building blocks of proteins 26
- The differing chemical and physical properties of amino acids are due to their side chains 27

Amino acids are covalently linked together in the protein chain by peptide bonds 29
Secondary structure of proteins is made up of α-helices and β-strands 33
Several different types of β-sheet are found in protein structures 35
Turns, hairpins and loops connect helices and strands 36

#### 2.2 Implication for Bioinformatics
- Certain amino acids prefer a particular structural unit 37
- Evolution has aided sequence analysis 38
- Visualization and computer manipulation of protein structures 38

#### 2.3 Proteins Fold to Form Compact Structures
- The tertiary structure of a protein is defined by the path of the polypeptide chain 40
- The stable folded state of a protein represents a state of low energy 41
- Many proteins are formed of multiple subunits 42

Summary 43
Further Reading 44

### Chapter 3 Dealing with Databases

#### 3.1 The Structure of Databases
- Flat-file databases store data as text files 46
- Relational databases are widely used for storing biological information 48
- XML has the flexibility to define bespoke data classifications 49
- Many other database structures are used for biological data 50
- Databases can be accessed locally or online and often link to each other 51

#### 3.2 Types of Database
- There's more to databases than just data 52
- Primary and derived data 53
- How we define and connect things is very important: Ontologies 54

#### 3.3 Looking for Databases
- Sequence databases 55
- Microarray databases 58
3.4 Data Quality
Nonredundancy is especially important for some applications of sequence databases. Automated methods can be used to check for data consistency. Initial analysis and annotation is usually automated. Human intervention is often required to produce the highest quality annotation. The importance of updating databases and entry identifier and version numbers.

Summary
Further Reading

Part 2 Sequence Alignments
APPLICATIONS CHAPTER
Chapter 4 Producing and Analyzing Sequence Alignments

4.1 Principles of Sequence Alignment
Alignment is the task of locating equivalent regions of two or more sequences to maximize their similarity. Alignment can reveal homology between sequences. It is easier to detect homology when comparing protein sequences than when comparing nucleic acid sequences.

4.2 Scoring Alignments
The quality of an alignment is measured by giving it a quantitative score. The simplest way of quantifying similarity between two sequences is percentage identity. The dot-plot gives a visual assessment of similarity based on identity. Genuine matches do not have to be identical. There is a minimum percentage identity that can be accepted as significant. There are many different ways of scoring an alignment.

4.3 Substitution Matrices
Substitution matrices are used to assign individual scores to aligned sequence positions. The PAM substitution matrices use substitution frequencies derived from sets of closely related protein sequences. The BLOSUM substitution matrices use mutation data from highly conserved local regions of sequence. The choice of substitution matrix depends on the problem to be solved.

4.4 Inserting Gaps
Gaps inserted in a sequence to maximize similarity require a scoring penalty. Dynamic programming algorithms can determine the optimal introduction of gaps.

4.5 Types of Alignment
Different kinds of alignments are useful in different circumstances. Multiple sequence alignments enable the simultaneous comparison of a set of similar sequences. Multiple alignments can be constructed by several different techniques. Multiple alignments can improve the accuracy of alignment for sequences of low similarity. ClustalW can make global multiple alignments of both DNA and protein sequences. Multiple alignments can be made by combining a series of local alignments. Alignment can be improved by incorporating additional information.

4.6 Searching Databases
Fast yet accurate search algorithms have been developed. FASTA is a fast database-search method based on matching short identical segments. BLAST is based on finding very similar short segments. Different versions of BLAST and FASTA are used for different problems. PSI-BLAST enables profile-based database searches. SSEARCH is a rigorous alignment method.

4.7 Searching with Nucleic Acid or Protein Sequences
DNA or RNA sequences can be used either directly or after translation. The quality of a database match has to be tested to ensure that it could not have arisen by chance. Choosing an appropriate E-value threshold helps to limit a database search. Low-complexity regions can complicate homology searches. Different databases can be used to solve particular problems.

4.8 Protein Sequence Motifs or Patterns
Creation of pattern databases requires expert knowledge. The BLOCKS database contains automatically compiled short blocks of conserved multiply aligned protein sequences.

4.9 Searching Using Motifs and Patterns
The PROSITE database can be searched for protein motifs and patterns.
The pattern-based program PHI-BLAST searches for both homology and matching motifs. Patterns can be generated from multiple sequences using PRATT. The PRINTS database consists of fingerprints representing sets of conserved motifs that describe a protein family. The Pfam database defines profiles of protein families.

### 4.10 Patterns and Protein Function

Searches can be made for particular functional sites in proteins. Sequence comparison is not the only way of analyzing protein sequences.

#### Summary

Further Reading

---

**THEORY CHAPTER**

**Chapter 5 Pairwise Sequence Alignment and Database Searching**

1. **Substitution Matrices and Scoring**
   - Alignment scores attempt to measure the likelihood of a common evolutionary ancestor.
   - The PAM (MDM) substitution scoring matrices were designed to trace the evolutionary origins of proteins.
   - The BLOSUM matrices were designed to find conserved regions of proteins.
   - Scoring matrices for nucleotide sequence alignment can be derived in similar ways.
   - The substitution scoring matrix used must be appropriate to the specific alignment problem.
   - Gaps are scored in a much more heuristic way than substitutions.

2. **Dynamic Programming Algorithms**
   - Optimal global alignments are produced using efficient variations of the Needleman–Wunsch algorithm.
   - Local and suboptimal alignments can be produced by making small modifications to the dynamic programming algorithm.
   - Time can be saved with a loss of rigor by not calculating the whole matrix.

3. **Indexing Techniques and Algorithmic Approximations**
   - Suffix trees locate the positions of repeats and unique sequences.
   - Hashing is an indexing technique that lists the starting positions of all k-tuples.
   - The FASTA algorithm uses hashing and chaining for fast database searching.

---

**Chapter 6 Patterns, Profiles, and Multiple Alignments**

1. **Profiles and Sequence Logos**
   - Position-specific scoring matrices are an extension of substitution scoring matrices.
   - Methods for overcoming a lack of data in deriving the values for a PSSM.
   - PSI-BLAST is a sequence database searching program.
   - Representing a profile as a logo.

2. **Profile Hidden Markov Models**
   - The basic structure of HMMs used in sequence alignment to profiles.
   - Estimating HMM parameters using aligned sequences.
   - Scoring a sequence against a profile HMM: The most probable path and the sum over all paths.
   - Estimating HMM parameters using unaligned sequences.

3. **Aligning Profiles**
   - Comparing two PSSMs by alignment.
   - Aligning profile HMMs.

4. **Multiple Sequence Alignments by Gradual Sequence Addition**
   - The order in which sequences are added is chosen based on the estimated likelihood of incorporating errors in the alignment.
   - Many different scoring schemes have been used in constructing multiple alignments.
6.5 Other Ways of Obtaining Multiple Alignments  
The multiple sequence alignment program DIALIGN aligns ungapped blocks  
The SAGA method of multiple alignment uses a genetic algorithm

6.6 Sequence Pattern Discovery  
Discovering patterns in a multiple alignment: eMOTIF and AACC  
Probabilistic searching for common patterns in sequences: Gibbs and MEME  
Searching for more general sequence patterns

Summary  
Further Reading

Part 3 Evolutionary Processes

Chapter 7 Recovering Evolutionary History  
7.1 The Structure and Interpretation of Phylogenetic Trees  
Phylogenetic trees reconstruct evolutionary relationships  
Tree topology can be described in several ways  
Consensus and condensed trees report the results of comparing tree topologies

7.2 Molecular Evolution and its Consequences  
Most related sequences have many positions that have mutated several times  
The rate of accepted mutation is usually not the same for all types of base substitution  
Different codon positions have different mutation rates  
Only orthologous genes should be used to construct species phylogenetic trees  
Major changes affecting large regions of the genome are surprisingly common

7.3 Phylogenetic Tree Reconstruction  
Small ribosomal subunit rRNA sequences are well suited to reconstructing the evolution of species  
The choice of the method for tree reconstruction depends on some extent on the size and quality of the dataset  
A model of evolution must be chosen to use with the method  
All phylogenetic analyses must start with an accurate multiple alignment

Phylogenetic analyses of a small dataset of 16S RNA sequence data  
Building a gene tree for a family of enzymes can help to identify how enzymatic functions evolved

Summary  
Further Reading

THEORY CHAPTER
Chapter 8 Building Phylogenetic Trees  
8.1 Evolutionary Models and the Calculation of Evolutionary Distance  
A simple but inaccurate measure of evolutionary distance is the p-distance  
The Poisson distance correction takes account of multiple mutations at the same site  
The Gamma distance correction takes account of mutation rate variation at different sequence positions  
The Jukes–Cantor model reproduces some basic features of the evolution of nucleotide sequences  
More complex models distinguish between the relative frequencies of different types of mutation  
There is a nucleotide bias in DNA sequences  
Models of protein-sequence evolution are closely related to the substitution matrices used for sequence alignment

8.2 Generating Single Phylogenetic Trees  
Clustering methods produce a phylogenetic tree based on evolutionary distances  
The UPGMA method assumes a constant molecular clock and produces an ultrametric tree  
The Fitch–Margoliash method produces an unrooted additive tree  
The neighbor-joining method is related to the concept of minimum evolution  
Stepwise addition and star-decomposition methods are usually used to generate starting trees for further exploration, not the final tree

8.3 Generating Multiple Tree Topologies  
The branch-and-bound method greatly improves the efficiency of exploring tree topology  
Optimization of tree topology can be achieved by making a series of small changes to an existing tree  
Finding the root gives a phylogenetic tree a direction in time

8.4 Evaluating Tree Topologies  
Functions based on evolutionary distances can be used to evaluate trees  
Unweighted parsimony methods look for the trees with the smallest number of mutations

Contents
Mutations can be weighted in different ways in the parsimony method 300
Trees can be evaluated using the maximum likelihood method 302
The quartet-puzzling method also involves maximum likelihood in the standard implementation 305
Bayesian methods can also be used to reconstruct phylogenetic trees 306

8.5 Assessing the Reliability of Tree Features and Comparing Trees 307
The long-branch attraction problem can arise even with perfect data and methodology 308
Tree topology can be tested by examining the interior branches 309
Tests have been proposed for comparing two or more alternative trees 310

Summary 311
Further Reading 312

Part 4 Genome Characteristics
APPLICATIONS CHAPTER
Chapter 9 Revealing Genome Features
9.1 Preliminary Examination of Genome Sequence 318
Whole genome sequences can be split up to simplify gene searches 319
Structural RNA genes and repeat sequences can be excluded from further analysis 319
Homology can be used to identify genes in both prokaryotic and eukaryotic genomes 322

9.2 Gene Prediction in Prokaryotic Genomes 322

9.3 Gene Prediction in Eukaryotic Genomes 323
Programs for predicting exons and introns use a variety of approaches 323
Gene predictions must preserve the correct reading frame 324
Some programs search for exons using only the query sequence and a model for exons 327
Some programs search for genes using only the query sequence and a gene model 332
Genes can be predicted using a gene model and sequence similarity 334
Genomes of related organisms can be used to improve gene prediction 336

9.4 Splice Site Detection 337
Splice sites can be detected independently by specialized programs 338

9.5 Prediction of Promoter Regions 338

THEORY CHAPTER
Chapter 10 Gene Detection and Genome Annotation
10.1 Detection of Functional RNA Molecules Using Decision Trees 361
Detection of tRNA genes using the tRNAscan algorithm 361
Detection of tRNA genes in eukaryotic genomes 362

10.2 Features Useful for Gene Detection in Prokaryotes 364
10.3 Algorithms for Gene Detection in Prokaryotes 368
GeneMark uses inhomogeneous Markov chains and codon statistics 368
GLIMMER uses interpolated Markov models of coding potential 371
ORPHEUS uses homology, codon statistics, and ribosome-binding sites 372
GeneMark.hmm uses explicit state duration hidden Markov models 373
EcoParse is an HMM gene model 376

10.4 Features Used in Eukaryotic Gene Detection 377
Differences between prokaryotic and eukaryotic genes 377
Introns, exons, and splice sites 379
Promoter sequences and binding sites for transcription factors 381
10.5 Predicting Eukaryotic Gene Signals
Detection of core promoter binding signals is a key element of some eukaryotic gene-prediction methods. A set of models has been designed to locate the site of core promoter sequence signals. Predicting promoter regions from general sequence properties can reduce the numbers of false-positive results. Predicting eukaryotic transcription and translation start sites. Translation and transcription stop signals complete the gene definition.

10.6 Predicting Exon/Intron Structure
Exons can be identified using general sequence properties. Splice-site prediction. Splice sites can be predicted by sequence patterns combined with base statistics. GenScan uses a combination of weight matrices and decision trees to locate splice sites. GeneSplicer predicts splice sites using first-order Markov chains. NetPlantGene uses neural networks with intron and exon predictions to predict splice sites. Other splicing features may yet be exploited for splice-site prediction. Specific methods exist to identify initial and terminal exons. Exons can be defined by searching databases for homologous regions.

10.7 Complete Eukaryotic Gene Models

10.8 Beyond the Prediction of Individual Genes

Summary
Further Reading

Part 5 Secondary Structures
APPLICATIONS CHAPTER
Chapter 11 Obtaining Secondary Structure from Sequence

11.1 Types of Prediction Methods
Statistical methods are based on rules that give the probability that a residue will form part of a particular secondary structure. Nearest-neighbor methods are statistical methods that incorporate additional information about protein structure. Machine-learning approaches to secondary structure prediction mainly make use of neural networks and HMM methods.

11.2 Training and Test Databases
There are several ways to define protein secondary structures.

11.3 Assessing the Accuracy of Prediction Programs
Q3 measures the accuracy of individual residue assignments. Secondary structure predictions should not be expected to reach 100% residue accuracy. The Sov value measures the prediction accuracy for whole elements. CAFASP/CASP: Unbiased and readily available protein prediction assessments.

11.4 Statistical and Knowledge-Based Methods
The GOR method uses an information theory approach. The program Zpred includes multiple alignment of homologous sequences and residue conservation information. There is an overall increase in prediction accuracy using multiple sequence information. The nearest-neighbor method: The use of multiple nonhomologous sequences. PREDATOR is a combined statistical and knowledge-based program that includes the nearest-neighbor approach.

11.5 Neural Network Methods of Secondary Structure Prediction
Assessing the reliability of neural net predictions. Several examples of Web-based neural network secondary structure prediction programs. PROF: Protein forecasting. PSIPRED. Jnet: Using several alternative representations of the sequence alignment.

11.6 Some Secondary Structures Require Specialized Prediction Methods
Transmembrane proteins. Quantifying the preference for a membrane environment.

11.7 Prediction of Transmembrane Protein Structure
Multi-helix membrane proteins. A selection of prediction programs to predict transmembrane helices.
12.4 Neural Networks Have Been Employed Successfully for Secondary Structure Prediction
Layered feed-forward neural networks can transform a sequence into a structural prediction
Inclusion of information on homologous sequences improves neural network accuracy
More complex neural nets have been applied to predict secondary and other structural features

12.5 Hidden Markov Models Have Been Applied to Structure Prediction
HMM methods have been found especially effective for transmembrane proteins
Nonmembrane protein secondary structures can also be successfully predicted with HMMs

12.6 General Data Classification Techniques Can Predict Structural Features
Support vector machines have been successfully used for protein structure prediction
Discriminants, SOMs, and other methods have also been used

11.8 Coiled-coil Structures
The COILS prediction program
PAIRCOIL and MULTICOIL are an extension of the COILS algorithm
Zipping the Leucine zipper: A specialized coiled coil

11.9 RNA Secondary Structure Prediction
Summary
Further Reading

THEORY CHAPTER
Chapter 12 Predicting Secondary Structures
12.1 Defining Secondary Structure and Prediction
Accuracy
The definitions used for automatic protein secondary structure assignment do not give identical results
There are several different measures of the accuracy of secondary structure prediction

12.2 Secondary Structure Prediction Based on Residue Propensities
Each structural state has an amino acid preference which can be assigned as a residue propensity
The simplest prediction methods are based on the average residue propensity over a sequence window
Residue propensities are modulated by nearby sequence
Predictions can be significantly improved by including information from homologous sequences

12.3 The Nearest-Neighbor Methods are Based on Sequence Segment Similarity
Short segments of similar sequence are found to have similar structure
Several sequence similarity measures have been used to identify nearest-neighbor segments
A weighted average of the nearest-neighbor segment structures is used to make the prediction
A nearest-neighbor method has been developed to predict regions with a high potential to misfold

12.4 Neural Networks Have Been Employed Successfully for Secondary Structure Prediction
Layered feed-forward neural networks can transform a sequence into a structural prediction
Inclusion of information on homologous sequences improves neural network accuracy
More complex neural nets have been applied to predict secondary and other structural features

12.5 Hidden Markov Models Have Been Applied to Structure Prediction
HMM methods have been found especially effective for transmembrane proteins
Nonmembrane protein secondary structures can also be successfully predicted with HMMs

12.6 General Data Classification Techniques Can Predict Structural Features
Support vector machines have been successfully used for protein structure prediction
Discriminants, SOMs, and other methods have also been used

11.9 RNA Secondary Structure Prediction
Summary
Further Reading

Part 6 Tertiary Structures
APPLICATIONS CHAPTER
Chapter 13 Modeling Protein Structure
13.1 Potential Energy Functions and Force Fields
The conformation of a protein can be visualized in terms of a potential energy surface
Conformational energies can be described by simple mathematical functions
Similar force fields can be used to represent conformational energies in the presence of averaged environments
Potential energy functions can be used to assess a modeled structure
Energy minimization can be used to refine a modeled structure and identify local energy minima
Molecular dynamics and simulated annealing are used to find global energy minima

13.2 Obtaining a Structure by Threading
The prediction of protein folds in the absence of known structural homologs
Libraries or databases of nonredundant protein folds are used in threading
Two distinct types of scoring schemes have been used in threading methods
Dynamic programming methods can identify optimal alignments of target sequences and structural folds
Several methods are available to assess the confidence to be put on the fold prediction. The C2-like domain from the Dictyostelia: A practical example of threading.

13.3 Principles of Homology Modeling
Closely related target and template sequences give better models. Significant sequence identity depends on the length of the sequence. Homology modeling has been automated to deal with the numbers of sequences that can now be modeled. Model building is based on a number of assumptions.

13.4 Steps in Homology Modeling
Structural homologs to the target protein are found in the PDB. Accurate alignment of target and template sequences is essential for successful modeling. The structurally conserved regions of a protein are modeled first. The modeled core is checked for misfits before proceeding to the next stage. Sequence realignment and remodeling may improve the structure. Insertions and deletions are usually modeled as loops. Nonidentical amino acid side chains are modeled mainly by using rotamer libraries. Energy minimization is used to relieve structural errors. Molecular dynamics can be used to explore possible conformations for mobile loops. Models need to be checked for accuracy. How far can homology models be trusted?

13.5 Automated Homology Modeling
The program MODELLER models by satisfying protein structure constraints. COMPOSER uses fragment-based modeling to automatically generate a model. Automated methods available on the Web for comparative modeling. Assessment of structure prediction.

13.6 Homology Modeling of PI3 Kinase p110α
Swiss-Pdb Viewer can be used for manual or semi-manual modeling. Alignment, core modeling, and side-chain modeling are carried out all in one. The loops are modeled from a database of possible structures. Energy minimization and quality inspection can be carried out within Swiss-Pdb Viewer.

MolIDE is a downloadable semi-automatic modeling package. Automated modeling on the Web illustrated with p110α kinase. Modeling a functionally related but sequentially dissimilar protein: mTOR. Generating a multidomain three-dimensional structure from sequence.

Summary
Further Reading

APPLICATIONS CHAPTER
Chapter 14 Analyzing Structure-Function Relationships

14.1 Functional Conservation
Functional regions are usually structurally conserved. Similar biochemical function can be found in proteins with different folds. Fold libraries identify structurally similar proteins regardless of function.

14.2 Structure Comparison Methods
Finding domains in proteins aids structure comparison. Structural comparisons can reveal conserved functional elements not discernible from a sequence comparison. The CE method builds up a structural alignment from pairs of aligned protein segments. The Vector Alignment Search Tool (VAST) aligns secondary structural elements. DALI identifies structure superposition without maintaining segment order. FATCAT introduces rotations between rigid segments.

14.3 Finding Binding Sites
Highly conserved, strongly charged, or hydrophobic surface areas may indicate interaction sites. Searching for protein–protein interactions using surface properties. Surface calculations highlight clefts or holes in a protein that may serve as binding sites. Looking at residue conservation can identify binding sites.

14.4 Docking Methods and Programs
Simple docking procedures can be used when the structure of a homologous protein bound to a ligand analog is known. Specialized docking programs will automatically dock a ligand to a structure.
Scoring functions are used to identify the most likely docked ligand.

The DOCK program is a semirigid-body method that analyzes shape and chemical complementarity of ligand and binding site.

Fragment docking identifies potential substrates by predicting types of atoms and functional groups in the binding area.

GOLD is a flexible docking program, which utilizes a genetic algorithm.

The water molecules in binding sites should also be considered.

Summary

Further Reading

Part 7 Cells and Organisms

Chapter 15 Proteome and Gene Expression Analysis

15.1 Analysis of Large-scale Gene Expression

The expression of large numbers of different genes can be measured simultaneously by DNA microarrays.

Gene expression microarrays are mainly used to detect differences in gene expression in different conditions.

Serial analysis of gene expression (SAGE) is also used to study global patterns of gene expression.

Digital differential display uses bioinformatics and statistics to detect differential gene expression in different tissues.

Facilitating the integration of data from different places and experiments.

The simplest method of analyzing gene expression microarray data is hierarchical cluster analysis.

Techniques based on self-organizing maps can be used for analyzing microarray data.

Self-organizing tree algorithms (SOTAs) cluster from the top down by successive subdivision of clusters.

Clustered gene expression data can be used as a tool for further research.

15.2 Analysis of Large-scale Protein Expression

Two-dimensional gel electrophoresis is a method for separating the individual proteins in a cell.

Measuring the expression levels shown in 2D gels.

Differences in protein expression levels between different samples can be detected by 2D gels.

Clustering methods are used to identify protein spots with similar expression patterns.

Principal component analysis (PCA) is an alternative to clustering for analyzing microarray and 2D gel data.

The changes in a set of protein spots can be tracked over a number of different samples.

Databases and online tools are available to aid the interpretation of 2D gel data.

Protein microarrays allow the simultaneous detection of the presence or activity of large numbers of different proteins.

Mass spectrometry can be used to identify the proteins separated and purified by 2D gel electrophoresis or other means.

Protein-identification programs for mass spectrometry are freely available on the Web.

Mass spectrometry can be used to measure protein concentration.
The self-organizing tree algorithm (SOTA) determines the number of clusters required. Biclustering identifies a subset of similar expression level patterns occurring in a subset of the samples. The validity of clusters is determined by independent methods.

16.4 Statistical Analysis can Quantify the Significance of Observed Differential Expression

$t$-tests can be used to estimate the significance of the difference between two expression levels. Nonparametric tests are used to avoid making assumptions about the data sampling. Multiple testing of differential expression requires special techniques to control error rates.

16.5 Gene and Protein Expression Data Can be Used to Classify Samples

Many alternative methods have been proposed that can classify samples. Support vector machines are another form of supervised learning algorithms that can produce classifiers.

Summary

Further Reading

Chapter 17 Systems Biology

17.1 What is a System?

A system is more than the sum of its parts. A biological system is a living network. Databases are useful starting points in constructing a network. To construct a model more information is needed than a network. There are three possible approaches to constructing a model. Kinetic models are not the only way in systems biology.

17.2 Structure of the Model

Control circuits are an essential part of any biological system. The interactions in networks can be represented as simple differential equations.

17.3 Robustness of Biological Systems

Robustness is a distinct feature of complexity in biology. Modularity plays an important part in robustness. Redundancy in the system can provide robustness. Living systems can switch from one state to another by means of bistable switches.