DETERMINING THE STRUCTURE OF BIOLOGICAL MACROMOLECULES: THEORY, MODELS, AND DATA

Agnes Bolinska

Department of History and Philosophy of Science

University of Cambridge

OUTLINE

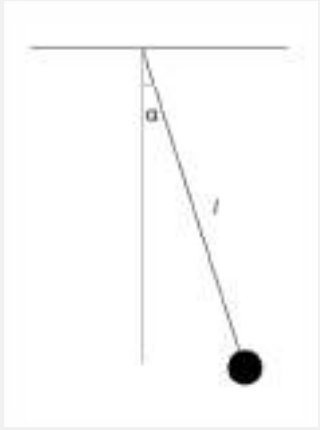
- What is a model? What is a theory?
  - The syntactic view of theories
  - The semantic view of theories
  - Models as mediators
- Challenges from DNA and protein structure determination
  - *Theories* as mediators?
  - Is the model/theory distinction important?

WHAT IS A MODEL?

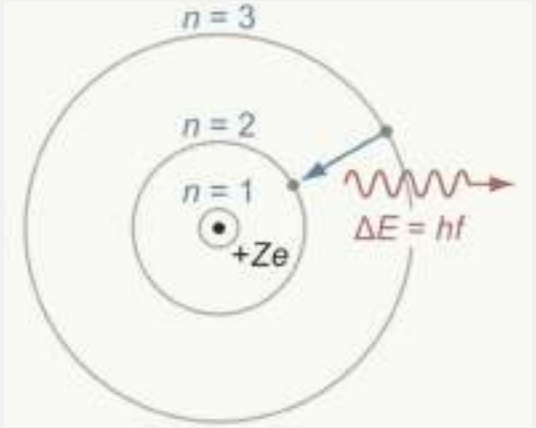
- Material models are physical things

WHAT IS A MODEL?

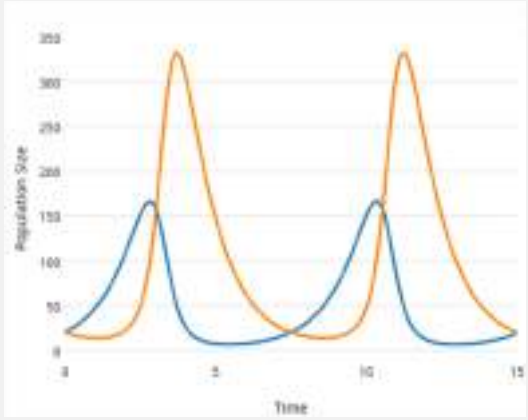
- But many models aren't physical things



Model of the simple pendulum



Bohr model of the atom



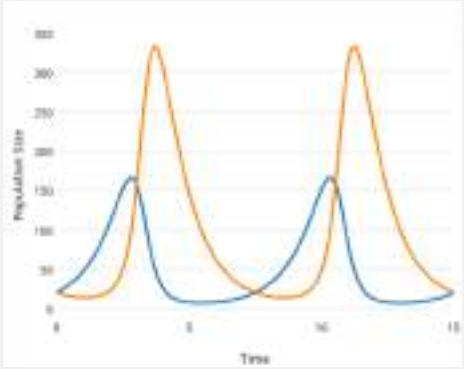
Lotka-Volterra model of predator-prey dynamics

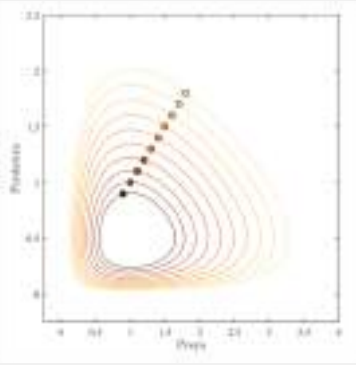
WHAT IS A MODEL?

- But many models aren't physical things
  - And they can be expressed in a variety of ways

$$\frac{dx}{dt} = \alpha x - \beta xy$$

$$\frac{dy}{dt} = \delta xy - \gamma y$$





WHAT IS A THEORY?

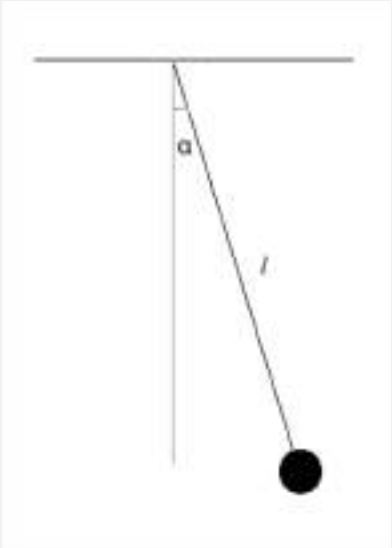
- The syntactic view of theories holds that theories are sets of sentences, relating different theoretical terms to one another by means of laws
- Meaning of theoretical terms was given in terms of their observational consequences via *correspondence rules*
  - e.g. ‘electron’ = ‘streak in a cloud chamber’
  - e.g. ‘mass’ is the result of a particular kind of measurement

MODELS AND THEORIES

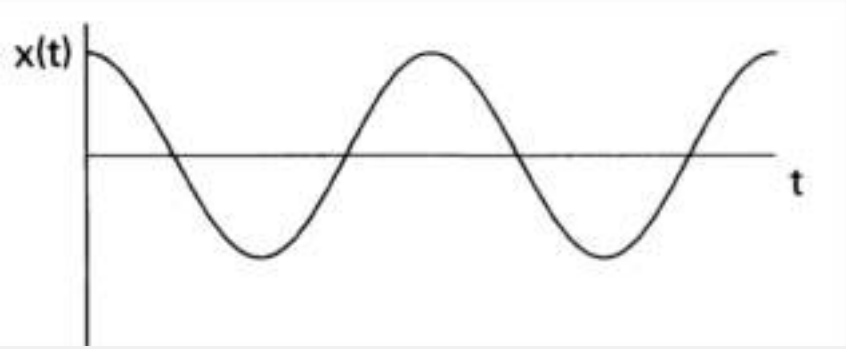
- Models can be "useful guides in suggesting theories," but are not "essential, even as psychological aids, and […] certainly not *logically* essential for a theory to be accepted as scientific" (Hesse 1966, p. 7)
- Models are not proper parts of scientific knowledge, which is the exclusive domain of theories

WHAT IS A THEORY?

- The semantic view of theories holds that theories are families of models
- Models constructed when laws of a theory are applied to a particular system
  - e.g. applying Newton's laws of motion to analyse the dynamics of a pendulum
- Models are non-linguistic entities
- Models, rather than correspondence rules, provide the interpretation of theories

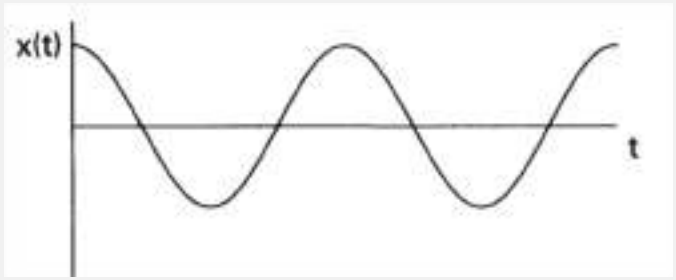


Theoretical model

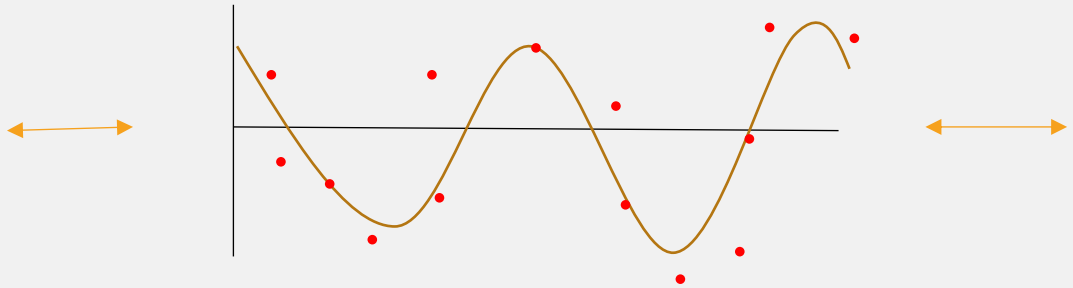


Real system

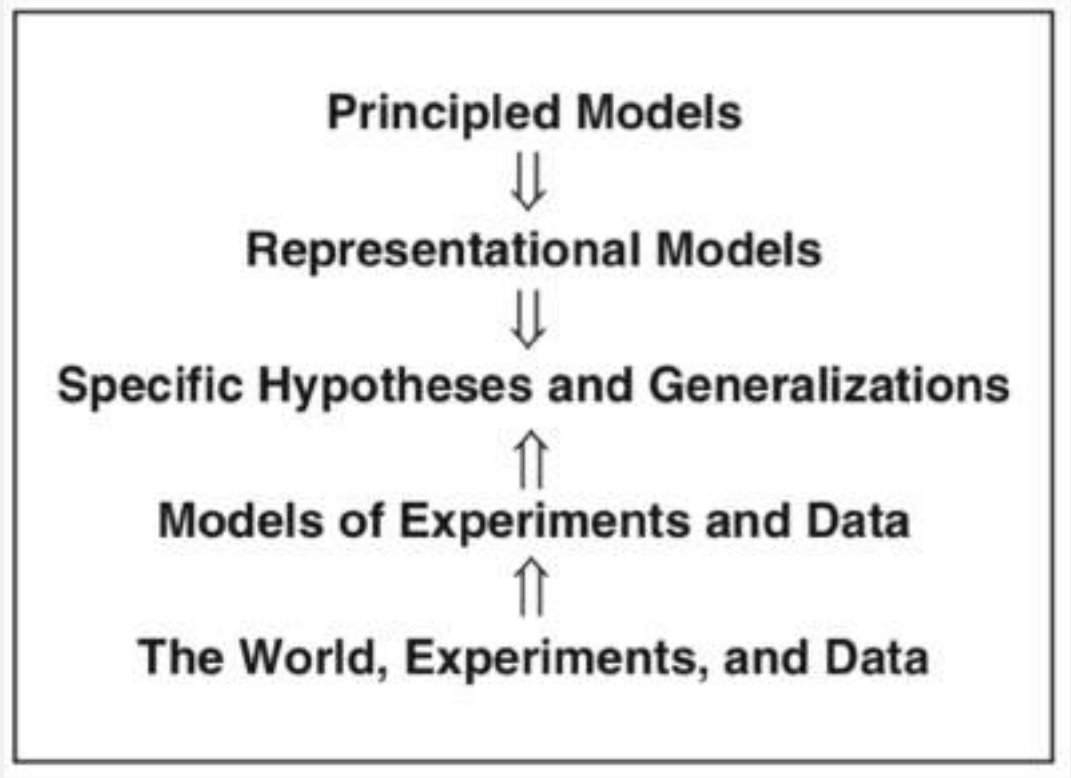
Theoretical model



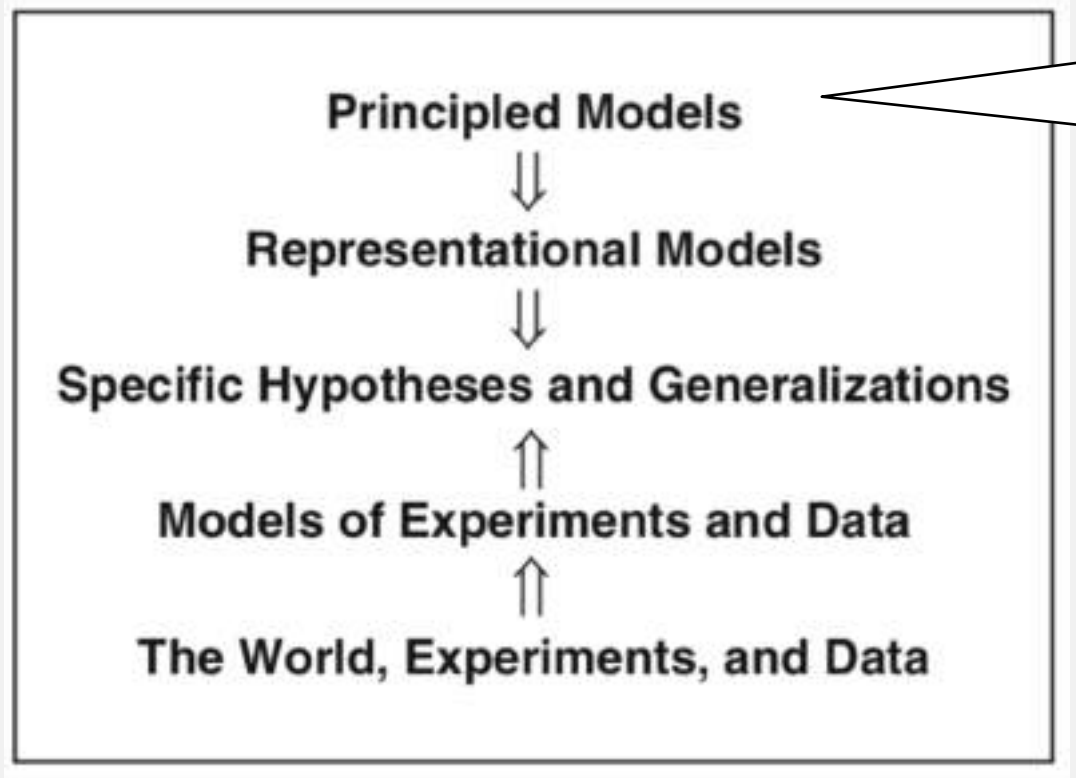
Model of data



Real system

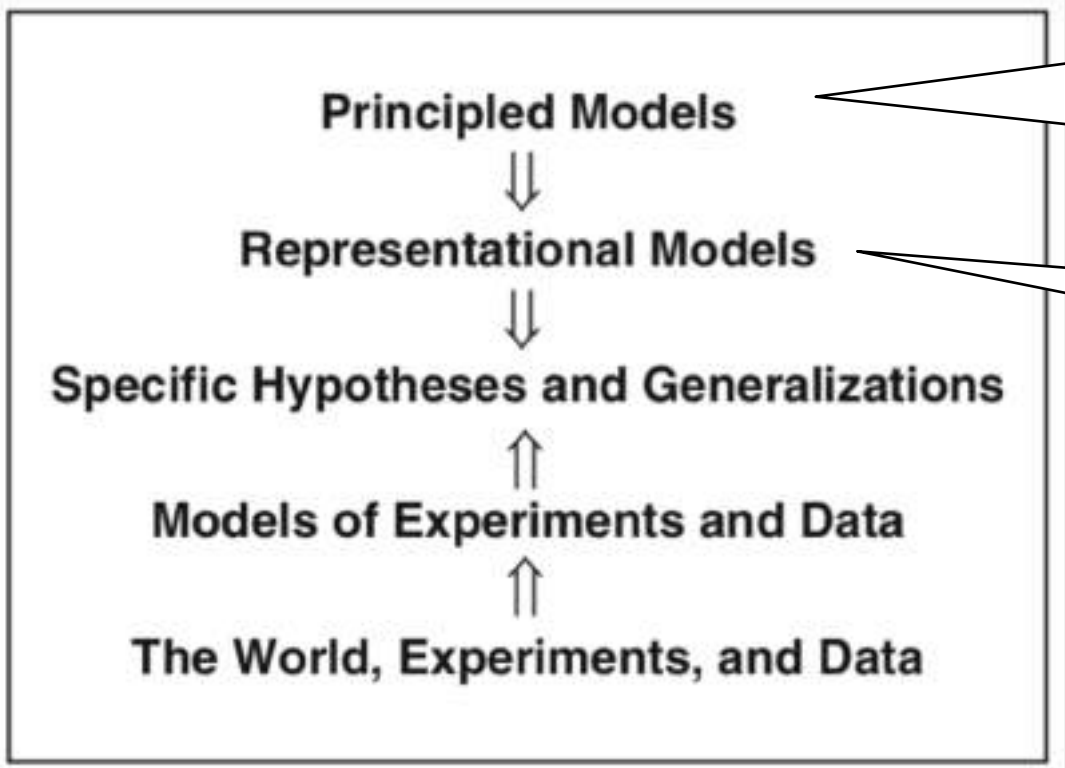


Giere 2010, p. 270



Giere 2010, p. 270

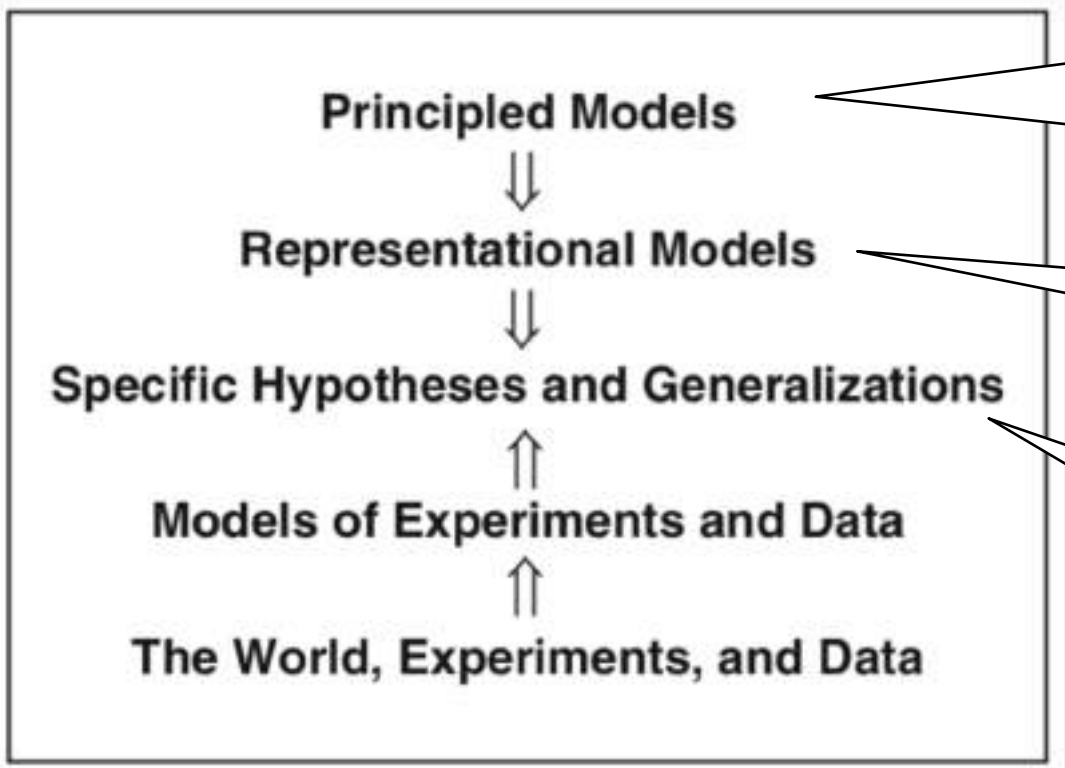
"statements that are often taken to constitute theories function to characterize the principled models," e.g. Newton's laws of motion (Giere 2010, p. 270)



"statements that are often taken to constitute theories function to characterize the principled models," e.g. Newton's laws of motion (Giere 2010, p. 270)

Model of the earth-moon system

Giere 2010, p. 270



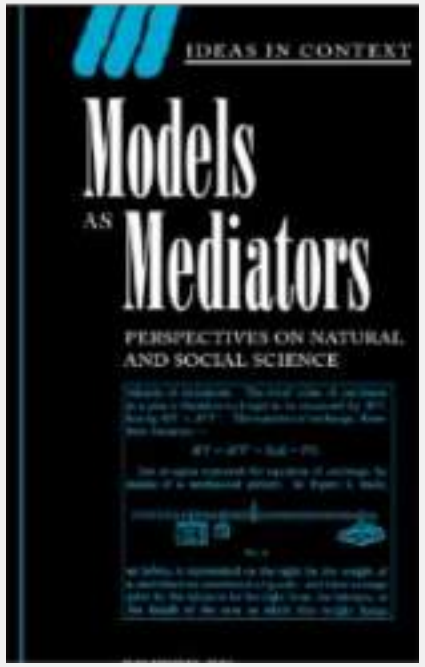
Giere 2010, p. 270

"statements that are often taken to constitute theories function to characterize the principled models," e.g. Newton's laws of motion (Giere 2010, p. 270)

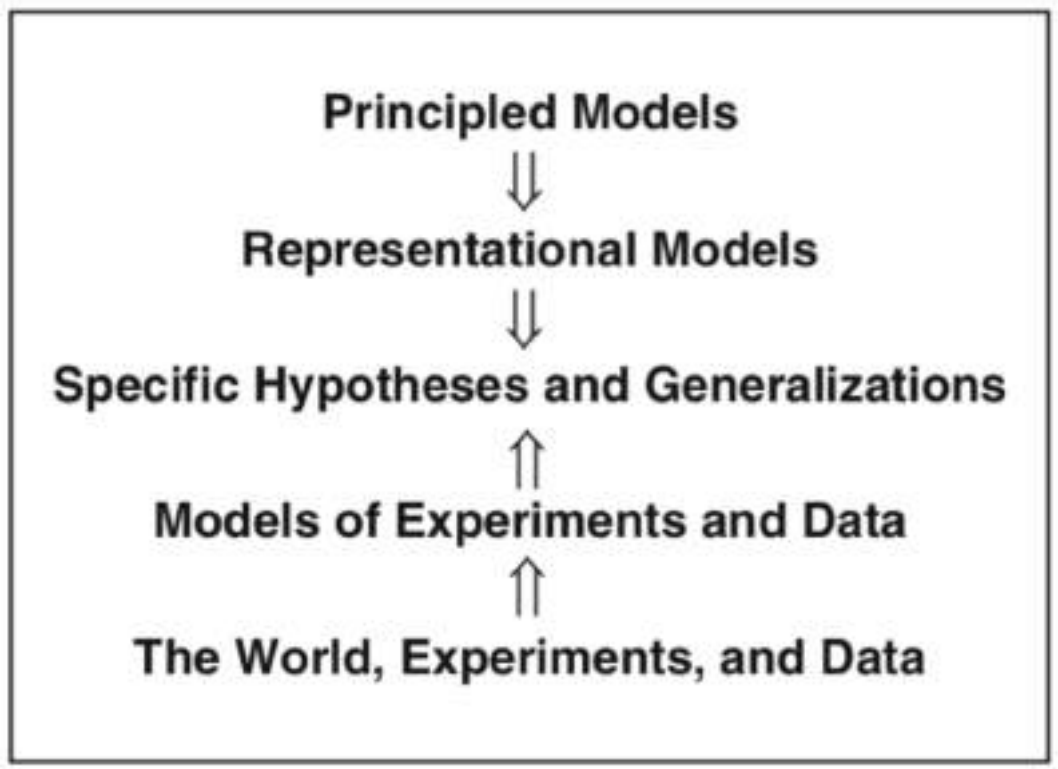
Model of the earth-moon system

"The positions and velocities of the earth and moon in the earth-moon system are very close to those of a two-particle Newtonian model with an inverse square force" (Giere 1988, p. 81)

MODELS AS MEDIATORS



- Models are 'autonomous agents': they are not constructed only on the basis of theory or data, but in a way that is partially dependent on both
- This autonomy enables them to be used as as 'instruments of investigation', enabling learning about theory and the world
- They are able to do this by representing some aspect of the theory, the world, or both
- Models always include elements of simplification and approximation that aren't dictated by the theory or data alone



Giere 2010, p. 270

"The crucial feature of partial independence is that models are *not* situated in the middle of an hierarchical structure between theory and the world. Because models typically include other elements, and model building proceeds in part independently of theory and data, we construe models as being outside the theory-world axis. It is this feature which enables them to mediate effectively between the two."

-Morrison and Morgan (1999, pp. 17-18)

"Mediating models always stand between theory and the physical world."

-Suárez (1999, p. 168)

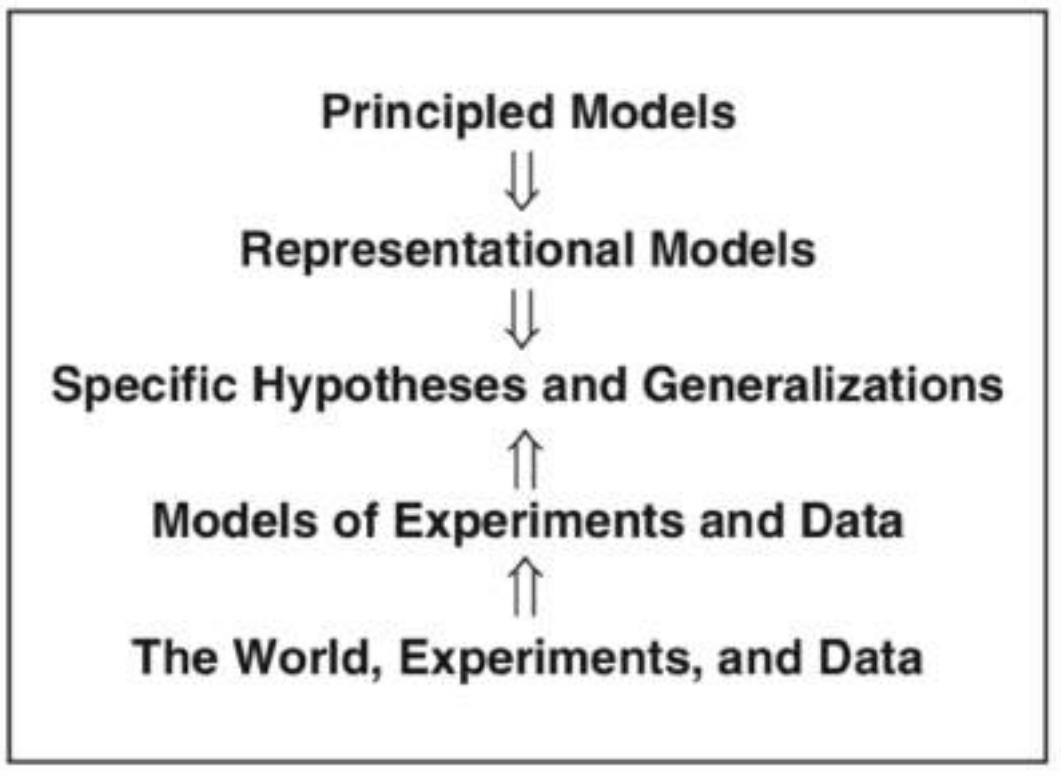
MODELS AS MEDIATORS

"My chief attack here is the … account of the relation of theory to model. This account gives us a kind of homunculus image of model creation: Theories have a belly-full of tiny already-formed models buried within them. It takes only the midwife of deduction to bring them forth."

-Cartwright, Shomar and Suárez (1995, p. 139)

MODELS AS MEDIATORS

"What is needed is the recognition of the independence from theory, in methods and aims, of the scientific activity we have come to call phenomenological model building." (*ibid.*, p. 148)



Giere 2010, p. 270

OUTLINE

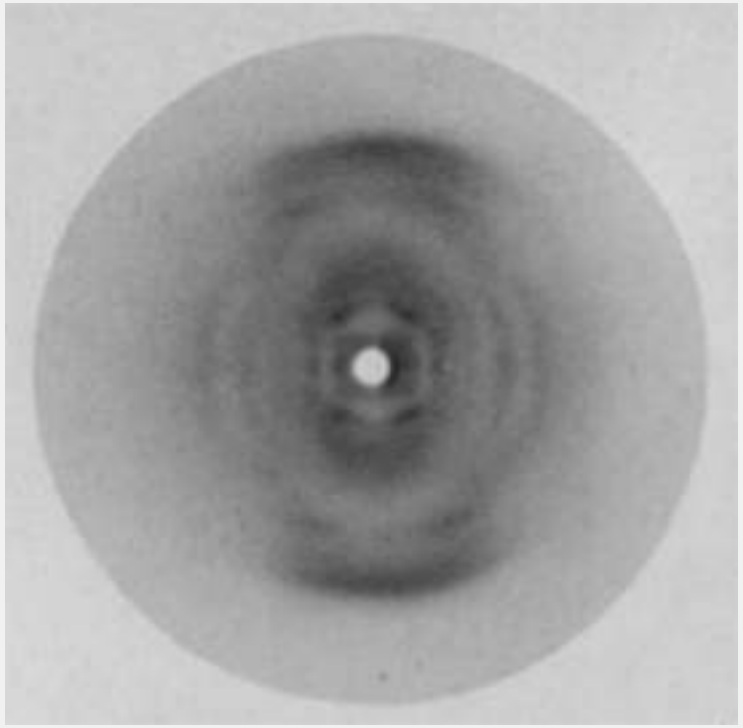
- What is a model? What is a theory?
  - The syntactic view of theories
  - The semantic view of theories
  - Models as mediators
- Challenges from DNA and protein structure determination
  - *Theories* as mediators?
  - Is the model/theory distinction important?

THEORIES AS MEDIATORS?

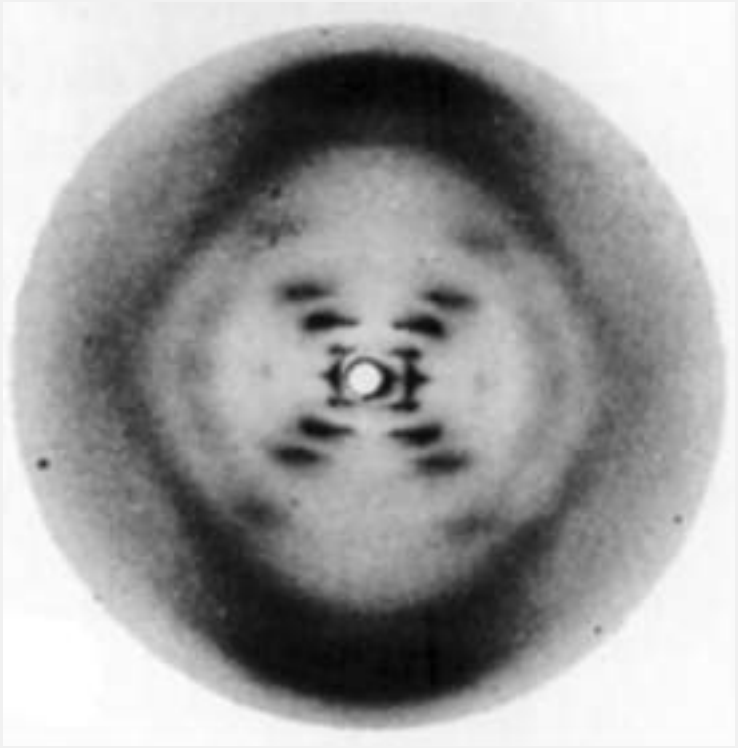
- Schindler (2008) argues that, in the determination of the DNA structure:
  - theory-driven model-building was successful; phenomenological model-building was not
  - *theory* mediated between model and data
  - a model was *deduced* from theory

HOW TO DETERMINE DNA STRUCTURE

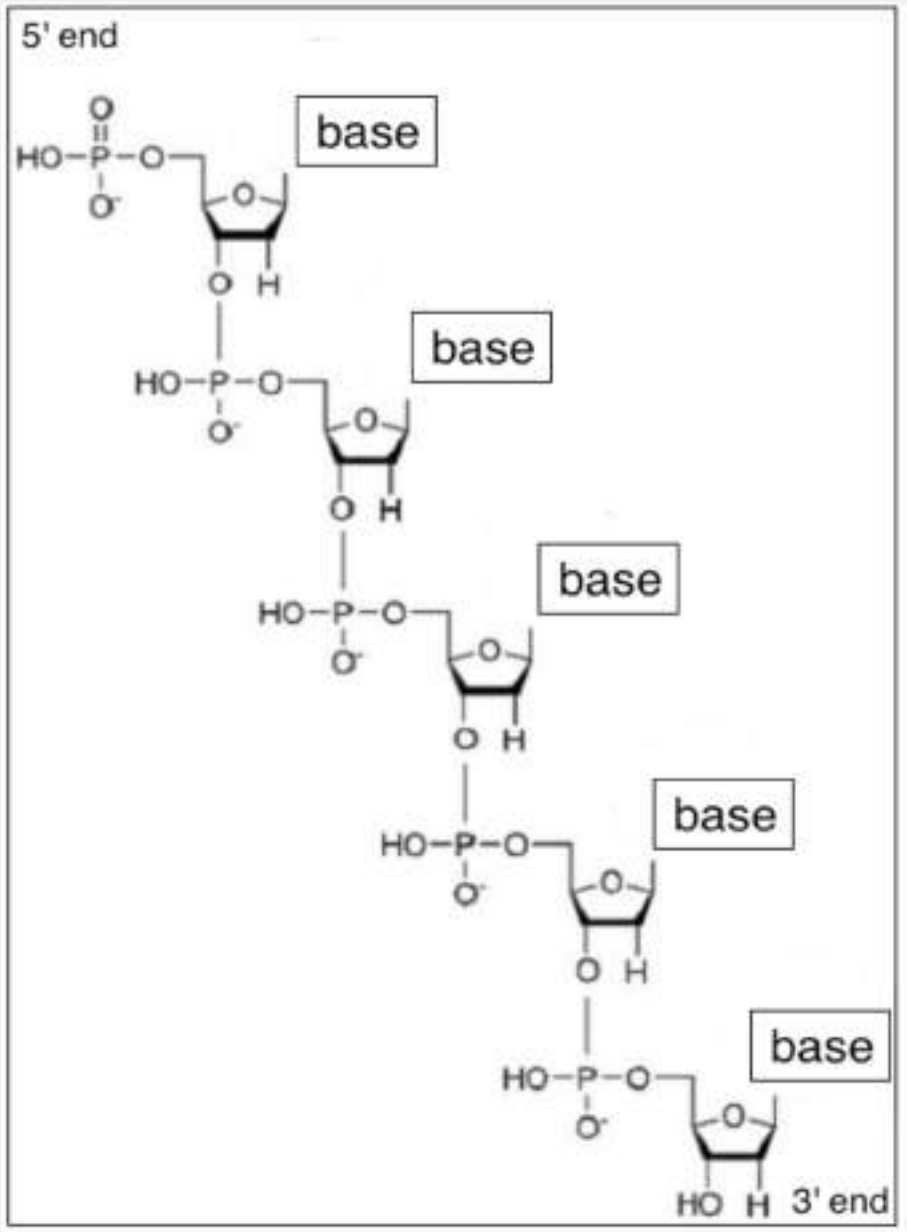
Option 1: infer the structure ‘bottom-up’ from X-ray diffraction photographs



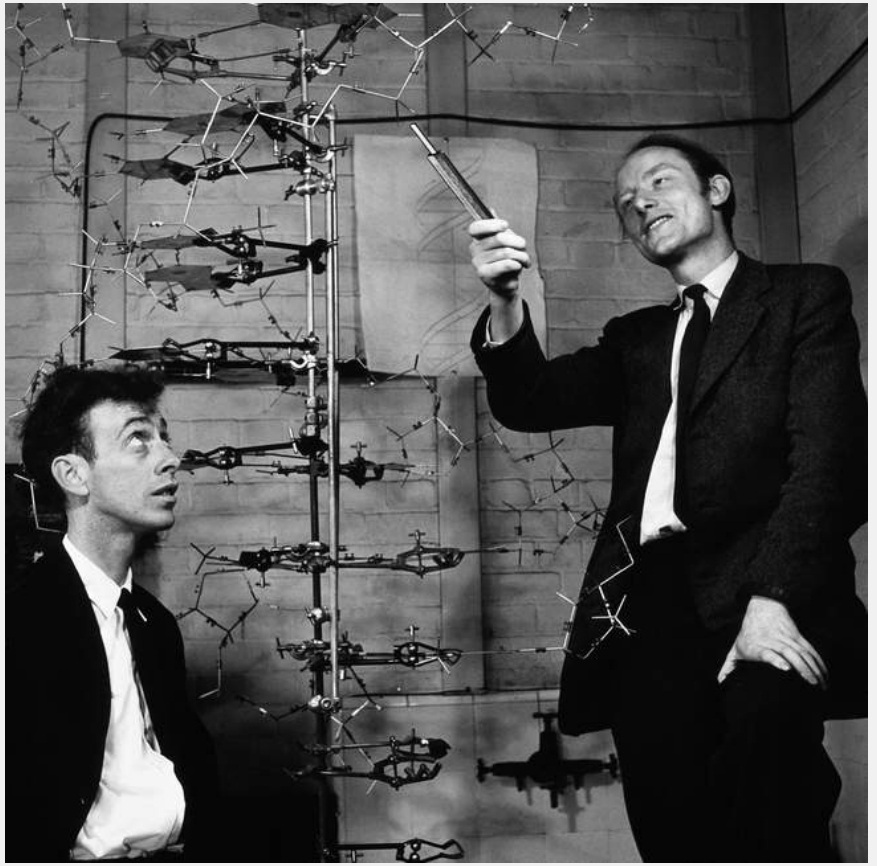
DNA. Reproduced from Astbury (1947)



The B form of DNA. Reproduced from Franklin & Gosling (1953)



Option 2: build a model ‘top-down’ based on known bond distances and angles



HOW TO DETERMINE DNA STRUCTURE

Watson and Crick built models from the ‘top down’

Franklin attempted to infer the structure from her photos from the ‘bottom up’

HOW TO DETERMINE DNA STRUCTURE

"And I felt repeatedly that Maurice was trying various ways to stimulate Rosalind into saying something about the structure, but she for her part would say, 'We are not going to speculate, we are going to wait, *we are going to let the spots on this photograph tell us what the structure is*.'"

-Gosling in Judson (1996, p. 127), quoted in Schindler (2008, p. 629); Schindler's emphasis

HOW TO DETERMINE DNA STRUCTURE

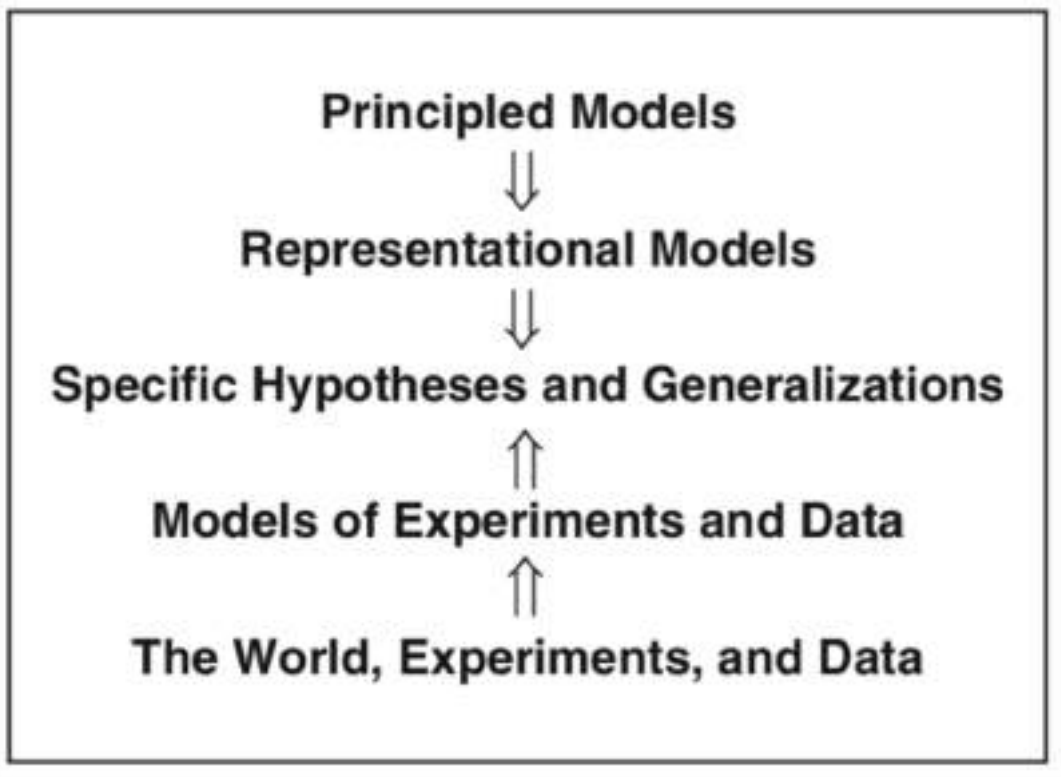
"I encouraged Bruce Fraser, in our lab, to try out his ideas in a model. Rosalind dismissed our excitement [about model-building] by saying that model-building is what you do *after* you have found the structure."

-Wilkins (2003, p. 160), quoted in Schindler (2008, p. 632); Schindler's emphasis

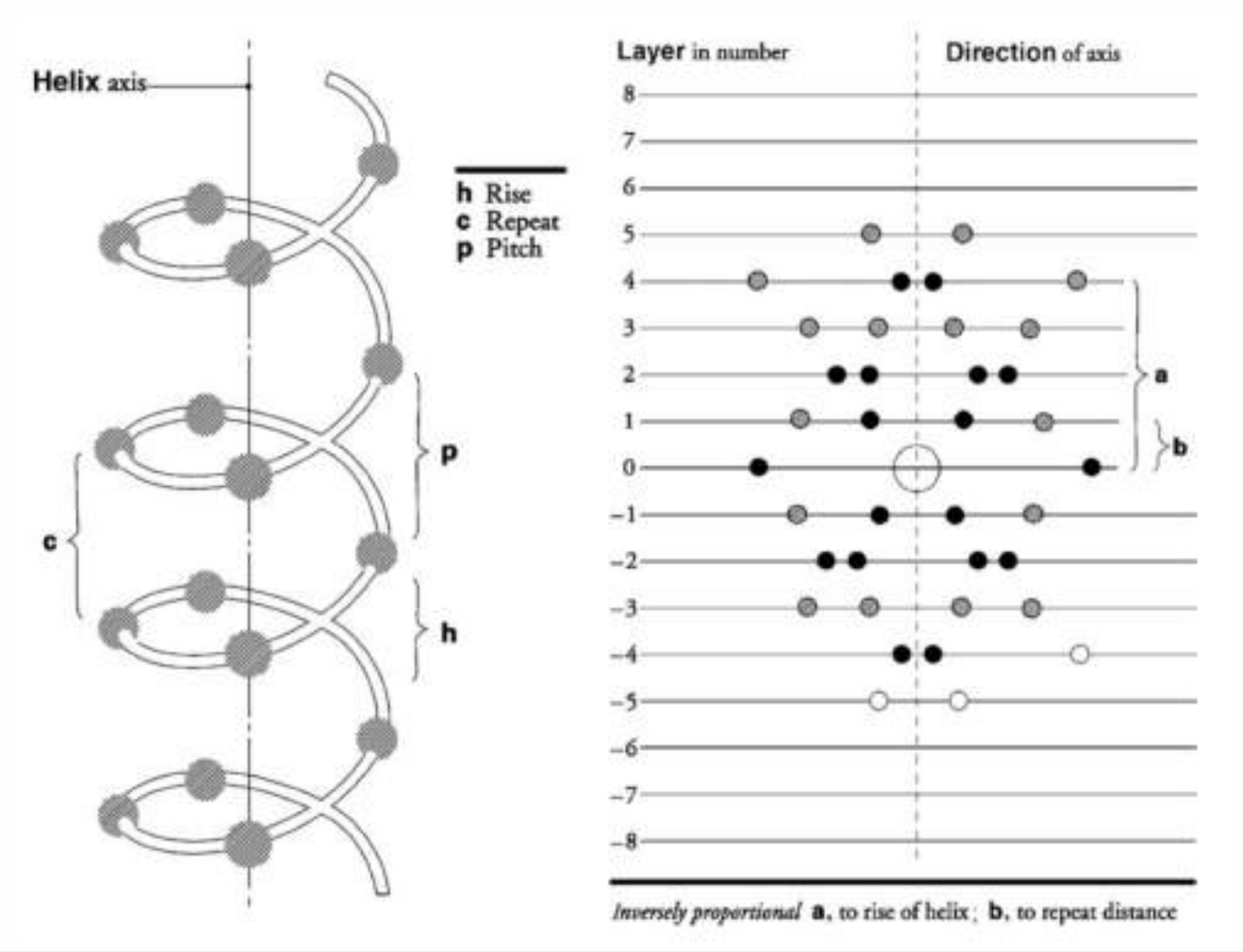
HOW TO DETERMINE DNA STRUCTURE

"[N]ot only is the method of building scale models an extremely powerful one, since it embodies a large amount of data which any successful model *must* include, but for structures of this type it *may well pay to build models without giving much attention to the experimental evidence*. [...] There is a case, in fact, for careful model building *independent of most of the experimental data*.

-Crick (1954b, p. 217), quoted in Schindler (2008, p. 633) Schindler's emphasis



Giere 2010, p. 270



Schematic representation of Cochrane, Crick, and Vand's (1952) helical diffraction theory (Schindler 2008, p. 643)

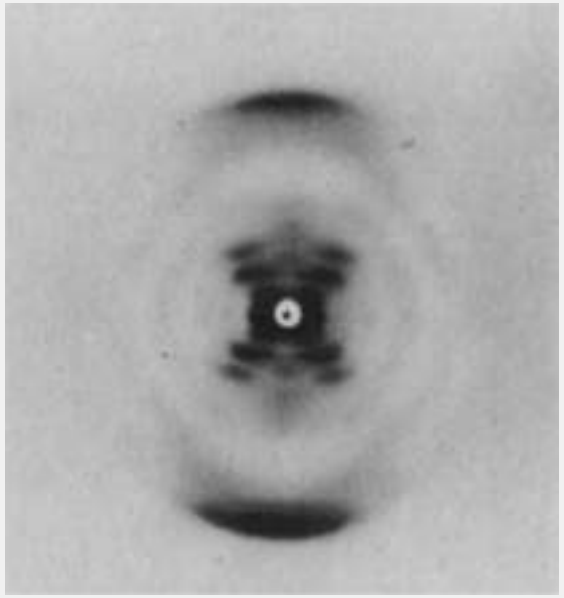
Schindler (2008): *theory* mediated between model and data

"*Armed with the appropriate theory* it is often possible to recognize the helical nature of a fiber structure *at a glance*, and sometimes to specify the main parameters of the helix and its subunits *with very little trouble* indeed."

-Crick and Kendrew (1957, p. 145), quoted in Schindler (2008, p. 642); Schindler's emphasis

HOW TO DETERMINE DNA STRUCTURE

- Schindler (2008): Cochrane, Crick, and Vand's helical diffraction theory was necessary to recognize that Franklin's B form indicated a helical structure
- 1951: before helical diffraction theory, Beighton's photo of the B-form was dismissed as a mixture, despite exhibiting an X-pattern



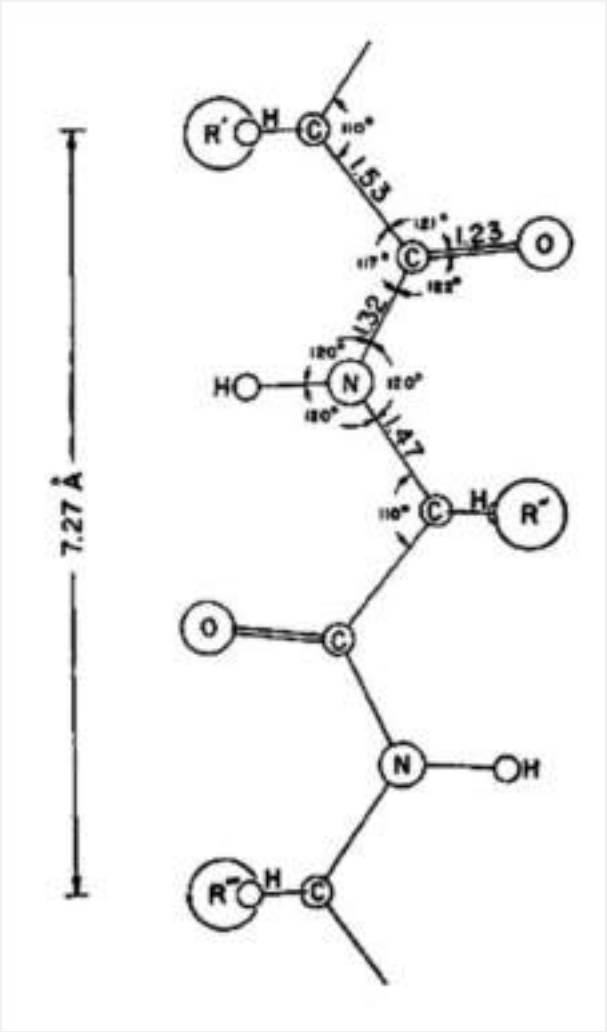
Beighton's B-form (Olby 1994)

"The [helical diffraction] theory thus established the *deductive* link between the structure of the model and the x-ray diffraction data. Only through this link from the top-down did the data gain the certainty and definiteness obvious from the quote above."

-Schindler (2008, p. 643)

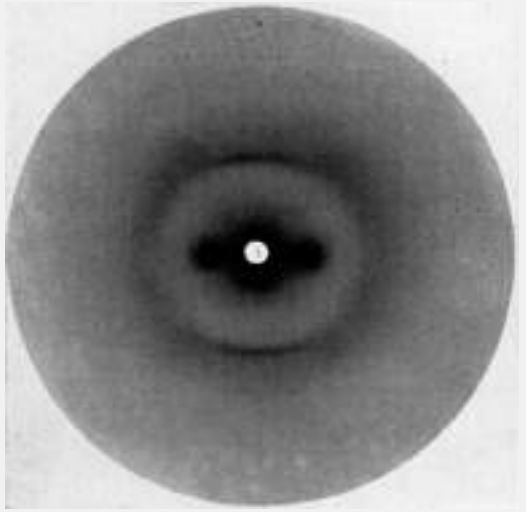
HOW TO DETERMINE DNA STRUCTURE

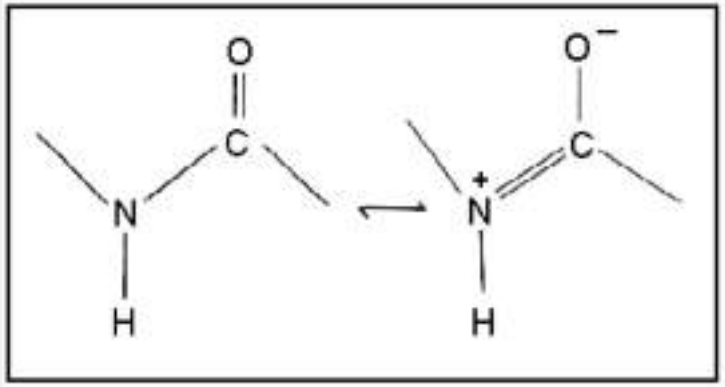
- Schindler (2008): Watson and Crick's approach was top-down, theory-driven; Franklin's was bottom-up, data-driven
- *Theory* established a *deductive* link between model and data in Watson and Crick's approach
- This runs against our usual intuitions about scientific practice and the idea that model-building is phenomenological
- So why did Watson and Crick advocate for an approach that ignored experimental evidence?



Pauling and Corey (1953)

- Astbury's photo seemed to indicate that protein would have a repeating subunit every 5.1 Å
- Pauling's α-helix rose 5.4 Å with each turn, *not* 5.1 Å
- 1950: Max Perutz, John Kendrew, and Sir Lawrence Bragg listed all structures compatible with X-ray diffraction photographs, chose the most likely candidate from amongst these
- Their proposed structure violated a stereochemical rule!





HOW TO DETERMINE DNA STRUCTURE

- Why did Watson and Crick advocate for an approach that ignored experimental evidence?
- Because they saw that ignoring experimental evidence can work, while paying too much attention to it can lead one astray
- But what was it about this approach that made it successful?
- Bolinska (2018): it adopted the heuristic likeliest to find the correct structure in the shortest period of time

HOW TO DETERMINE DNA STRUCTURE

Watson and Crick's approach was synthetic: they built the structure up from component parts

Franklin's was analytic: she attempted to derive structure from X-ray diffraction photographs

FORMULATING THE PROBLEM

- The problem of molecular structure determination:
  - Begins with a space of candidate structures, each a possible solution to the problem
  - The size of this space is reduced through the consideration of various *constraints* on structure, bits of information about bond types, lengths, and angles, and X-ray diffraction photographs of the molecule

HOW TO DETERMINE DNA STRUCTURE

Watson and Crick's approach was synthetic: they built the structure up from component parts

Franklin's was analytic: she attempted to derive structure from X-ray diffraction photographs

HOW TO DETERMINE DNA STRUCTURE

Watson and Crick's adopted a components-first heuristic

Franklin adopted a photos-first heuristic

HOW TO DETERMINE DNA STRUCTURE

- Bolinska (2018): all else being equal, the components-first heuristic was most likely to be successful because scientists adopting this heuristic could be more confident that they had eliminated candidate structures *correctly*
- Bond distances and angles were determined through repeated X-ray diffraction studies on small molecules
- "[E]xact and careful model building could embody constraints that the final answer had in any case to satisfy" (Crick 1988, p. 60)
- Is the theory/model distinction important?

IS THE THEORY/MODEL DISTINCTION IMPORTANT?

"[N]ot only is the method of building scale models an extremely powerful one, since it embodies a large amount of data which any successful model *must* include, but for structures of this type it *may well pay to build models without giving much attention to the experimental evidence*. [...] There is a case, in fact, for careful model building *independent of most of the experimental data*.

-Crick (1954b, p. 217), quoted in Schindler (2008, p. 633) Schindler's emphasis

IS THE THEORY/MODEL DISTINCTION IMPORTANT?

- Watson and Crick's model-building relied on experimental data
- Franklin's interpretation of X-ray diffraction photos relied on theory

SUMMARY

- On the semantic view, (representational/theoretical) models are deductively derived from theory
- The models-as-mediators view holds that models are neither derived exclusively from theory or from data, but are partially independent of both; their autonomy enables them to mediate between theory and data
- Schindler (2008) argues that, in the case of DNA structure determination, the model was deduced from theory
- And helical diffraction theory mediated between model and data
- Bolinska (2018) shows that the heuristics for determining DNA structure are not neatly decomposed into theory-driven and data-driven ones

QUESTIONS

- Is the link between data and model established by helical diffraction theory really a *deductive* one?
- Does the case of DNA undermine the models-as-mediators view, given the disparate ways in which "model" is understood?
  - Could Franklin be understood as constructing a phenomenological model, albeit not a physical one?
  - Even if not, should the models-as-mediators view be interpreted as applying to *every* instance of scientific practice?
- Is analyzing the theory/model relation a useful focus for philosophical work?

Thank you!

amb273@cam.ac.uk

References:

- Astbury, W. T. (1947). X-ray studies of nucleic acids. *Symposia Society for Experimental Biology*, *1*, 66–76.
- Bolinska, A. (2018) 'Synthetic versus analytic approaches to protein and DNA structure determination', *Biology and Philosophy* 33: 26.
- Cartwright, N., Suarez, M. and Shomar, T. (1995). The Tool Box of Science. Tools for the Building of Models with a Superconductivity Example, in W. Herfel, W. Krajewski, I. Niiniluoto and R. Wojcicki (eds), *Theories and Models in Scientific Process*, Poznan Studies in the Philosophy of Science and the Humanities, 44, Amsterdam: Rodopi, pp. 137–49.
- Crick, F. H. C. (1988). *What Mad Pursuit: A Personal View of Scientific Discovery*. New York: Basic Books.
- Franklin, R., & Gosling, R. (1953). The Structure of Sodium-Thymonucleate Fibres I. The Influence of water content. *Acta Crystallographica*, *6*.
- Giere, R. N. (2010). An agent-based conception of models and scientific representation. *Synthese*, *172*(2), 269–281.
- Hesse, M. (1966). *Models and Analogies in Science*. Notre Dame: University of Notre Dame Press.
- Morgan, M. S., & Morrison, M. (Eds.). (1999). *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.
- Olby, R. (1974). *The Path to the Double Helix: The Discovery of DNA*. London: MacMillan.
- Schindler, S. (2008). Model, Theory, and Evidence in the Discovery of the DNA Structure. *The British Journal for the Philosophy of Science*, *59*(4), 619–658.
- Suárez, M. (1999). The role of models in the application of scientific theories: epistemological implications. In Morgan, M. S., & Morrison, M. (Eds.). *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.