Examination of Protein Crystallography Applications

Kathleen Kerr Bragdon, Ph.D.
Supervisory Patent Examiner
Art Unit 1656

Overview

- What is protein crystallography subject matter?
- Basis for discussion:

  Trilateral Co-operation Biotechnology Project
- Statutory Subject Matter - 35 USC § 101
- Written Description and Enablement - 35 USC § 112, first paragraph
- Prior Art - 35 USC § 102 and 103

Protein Crystallography Subject Matter

- Protein crystallography

  The art of getting a protein to “sit still”

  …then taking a 3D “picture”

- What are protein crystals?

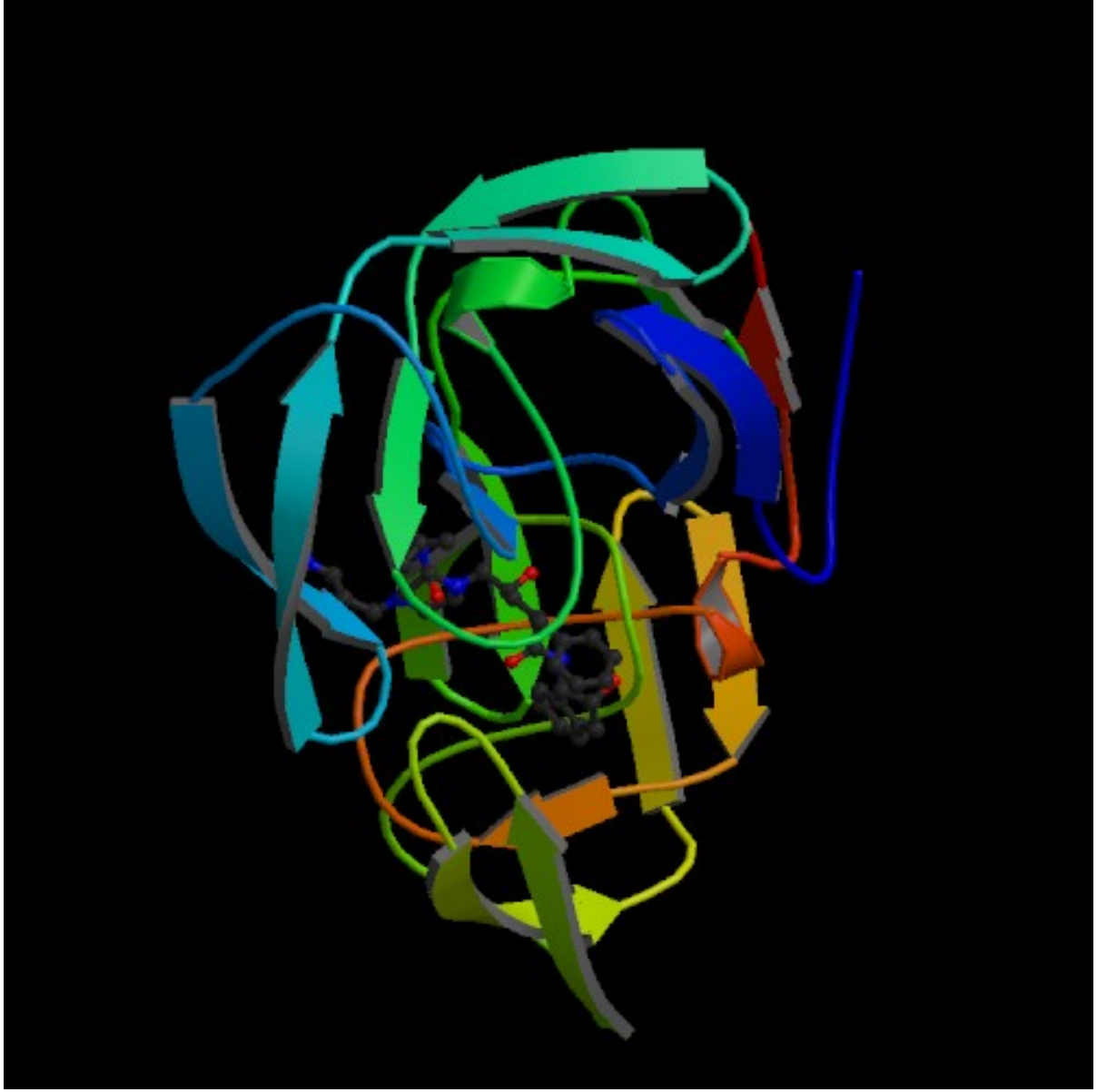
  Static, well-ordered arrays of protein molecules

- How are these “pictures” made?

  By projecting x-rays through the ordered protein arrays, collecting the constructively diffracted x-rays, and reconstructing a likely model of the protein’s 3D structure

Biotechnology in 3-Dimensions

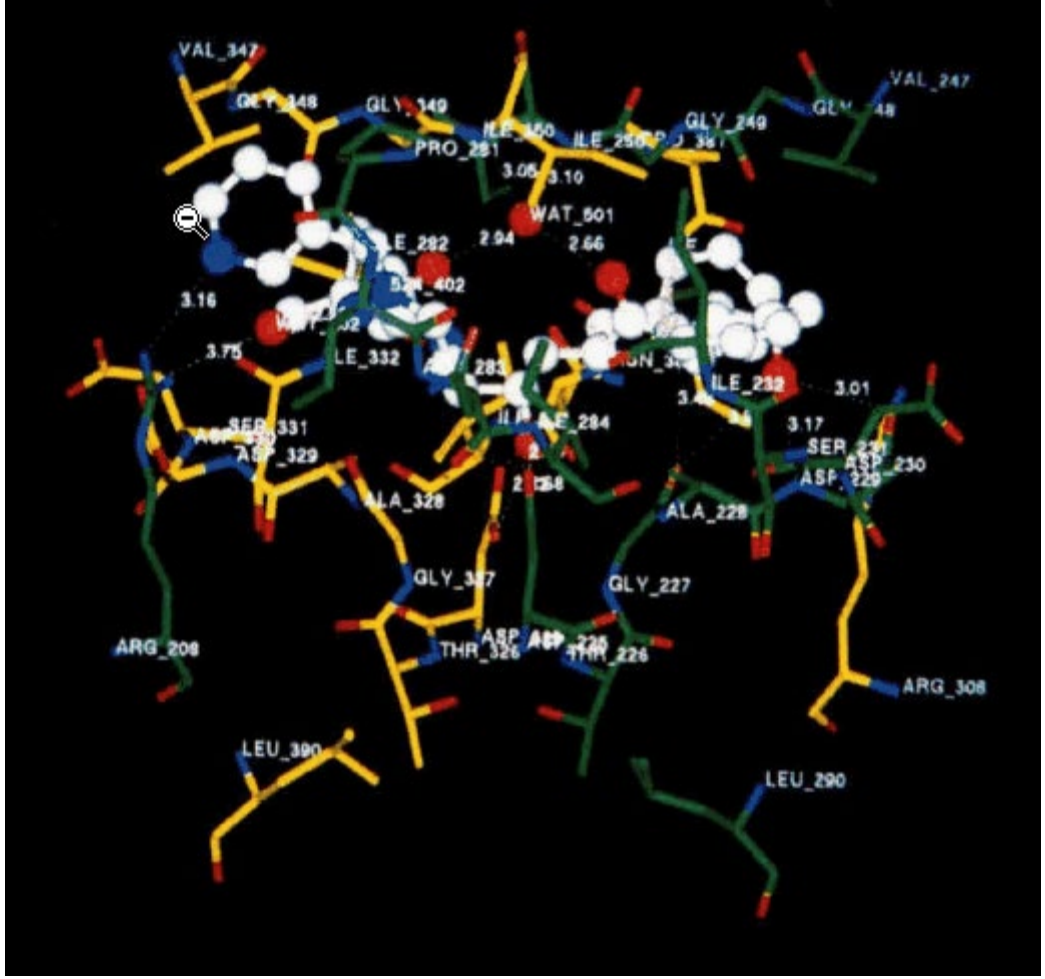
- Useful for identifying inhibitors & agonists



From Protein Data Bank (PDB) file **1HSG**
Crystal Structure at 1.9 A Resolution of HIV II Protease
J.Biol.Chem. v269 pp.26344-26348 , 1994

```
HEADER    HYDROLASE (ACID PROTEINASE)                 31-MAR-95   1HSG      1HSG   2
COMPND   2 MOLECULE: HIV-1 PROTEASE;                                        1HSG   4
COMPND   3 CHAIN: A, B;                                                     1HSG   5
SOURCE   2 ORGANISM_SCIENTIFIC: HUMAN IMMUNODEFICIENCY VIRUS TYPE 1;        1HSG  10
SOURCE   3 GENE: HIV-1 PROTEASE FROM THE NY5 ISOLATE;                       1HSG  11
EXPDTA    X-RAY DIFFRACTION                                                 1HSG  13
REMARK   2                                                                  1HSG  25
REMARK   2 RESOLUTION. 2.0  ANGSTROMS.                                      1HSG  26
REMARK   3   R VALUE                      0.166                             1HSG  31
REMARK   3   RMSD BOND DISTANCES          0.017  ANGSTROMS                  1HSG  32
REMARK   3   RMSD BOND ANGLES            1.9    DEGREES                    1HSG  33

SEQRES   1 A   99  PRO GLN ILE THR LEU TRP GLN ARG PRO LEU VAL THR ILE      1HSG  62

ATOM      1  N   PRO A   1      29.361  39.686   5.862  1.00 38.10      1HSG 107
ATOM      2  CA  PRO A   1      30.307  38.663   5.319  1.00 40.62      1HSG 108
ATOM      3  C   PRO A   1      29.760  38.071   4.022  1.00 42.64      1HSG 109
ATOM      4  O   PRO A   1      28.600  38.302   3.676  1.00 43.40      1HSG 110
ATOM      5  CB  PRO A   1      30.508  37.541   6.342  1.00 37.87      1HSG 111
ATOM      6  CG  PRO A   1      29.296  37.591   7.162  1.00 38.40      1HSG 112
ATOM      7  CD  PRO A   1      28.778  39.015   7.019  1.00 38.74      1HSG 113
ATOM      8  N   GLN A   2      30.607  37.334   3.305  1.00 41.76      1HSG 114
ATOM      9  CA  GLN A   2      30.158  36.492   2.199  1.00 41.30      1HSG 115
ATOM     10  C   GLN A   2      30.298  35.041   2.643  1.00 41.38      1HSG 116
ATOM     11  O   GLN A   2      31.401  34.494   2.763  1.00 43.09      1HSG 117
ATOM     12  CB  GLN A   2      30.970  36.738   0.926  1.00 40.81      1HSG 118
ATOM     13  CG  GLN A   2      30.625  35.783  -0.201  1.00 46.61      1HSG 119
ATOM     14  CD  GLN A   2      31.184  36.217  -1.549  1.00 50.36      1HSG 120
ATOM     15  OE1 GLN A   2      32.006  35.518  -2.156  1.00 53.89      1HSG 121
ATOM     16  NE2 GLN A   2      30.684  37.339  -2.061  1.00 51.46      1HSG 122
ATOM     17  N   ILE A   3      29.160  34.436   2.919  1.00 37.80      1HSG 123
ATOM     18  CA  ILE A   3      29.123  33.098   3.397  1.00 34.13      1HSG 124
ATOM     19  C   ILE A   3      28.968  32.155   2.198  1.00 33.19      1HSG 125
ATOM     20  O   ILE A   3      28.088  32.330   1.368  1.00 32.74      1HSG 126
```

```
HEADER    HYDROLASE (ACID PROTEINASE)                 31-MAR-95   1HSG       1HSG   2
COMPND    2 MOLECULE: HIV-1 PROTEASE;                                        1HSG   4
COMPND    3 CHAIN: A, B;                                                     1HSG   5
SOURCE    2 ORGANISM_SCIENTIFIC: HUMAN IMMUNODEFICIENCY VIRUS TYPE 1;        1HSG  10
SOURCE    3 GENE: HIV-1 PROTEASE FROM THE NY5 ISOLATE;                       1HSG  11
EXPDTA    X-RAY DIFFRACTION                                                  1HSG  13
REMARK    2                                                                  1HSG  25
REMARK    2 RESOLUTION. 2.0  ANGSTROMS.                                      1HSG  26
REMARK    3   R VALUE                      0.166                             1HSG  31
REMARK    3   RMSD BOND DISTANCES          0.017  ANGSTROMS                  1HSG  32
REMARK    3   RMSD BOND ANGLES             1.9    DEGREES                    1HSG  33

SEQRES    1 A   99  PRO GLN ILE THR LEU TRP GLN ARG PRO LEU VAL THR ILE      1HSG  62

ATOM      1  N   PRO A   1      29.361  39.686   5.862  1.00 38.10       1HSG 107
ATOM      2  CA  PRO A   1      30.307  38.663   5.319  1.00 40.62       1HSG 108
ATOM      3  C   PRO A   1      29.760  38.071   4.022  1.00 42.64       1HSG 109
ATOM      4  O   PRO A   1      28.600  38.302   3.676  1.00 43.40       1HSG 110
ATOM      5  CB  PRO A   1      30.508  37.541   6.342  1.00 37.87       1HSG 111
ATOM      6  CG  PRO A   1      29.296  37.591   7.162  1.00 38.40       1HSG 112
ATOM      7  CD  PRO A   1      28.778  39.015   7.019  1.00 38.74       1HSG 113
ATOM      8  N   GLN A   2      30.607  37.334   3.305  1.00 41.76       1HSG 114
ATOM      9  CA  GLN A   2      30.158  36.492   2.199  1.00 41.30       1HSG 115
ATOM     10  C   GLN A   2      30.298  35.041   2.643  1.00 41.38       1HSG 116
ATOM     11  O   GLN A   2      31.401  34.494   2.763  1.00 43.09       1HSG 117
ATOM     12  CB  GLN A   2      30.970  36.738   0.926  1.00 40.81       1HSG 118
ATOM     13  CG  GLN A   2      30.625  35.783  -0.201  1.00 46.61       1HSG 119
ATOM     14  CD  GLN A   2      31.184  36.217  -1.549  1.00 50.36       1HSG 120
ATOM     15  OE1 GLN A   2      32.006  35.518  -2.156  1.00 53.89       1HSG 121
ATOM     16  NE2 GLN A   2      30.684  37.339  -2.061  1.00 51.46       1HSG 122
ATOM     17  N   ILE A   3      29.160  34.436   2.919  1.00 37.80       1HSG 123
ATOM     18  CA  ILE A   3      29.123  33.098   3.397  1.00 34.13       1HSG 124
ATOM     19  C   ILE A   3      28.968  32.155   2.198  1.00 33.19       1HSG 125
ATOM     20  O   ILE A   3      28.088  32.330   1.368  1.00 32.74       1HSG 126
```

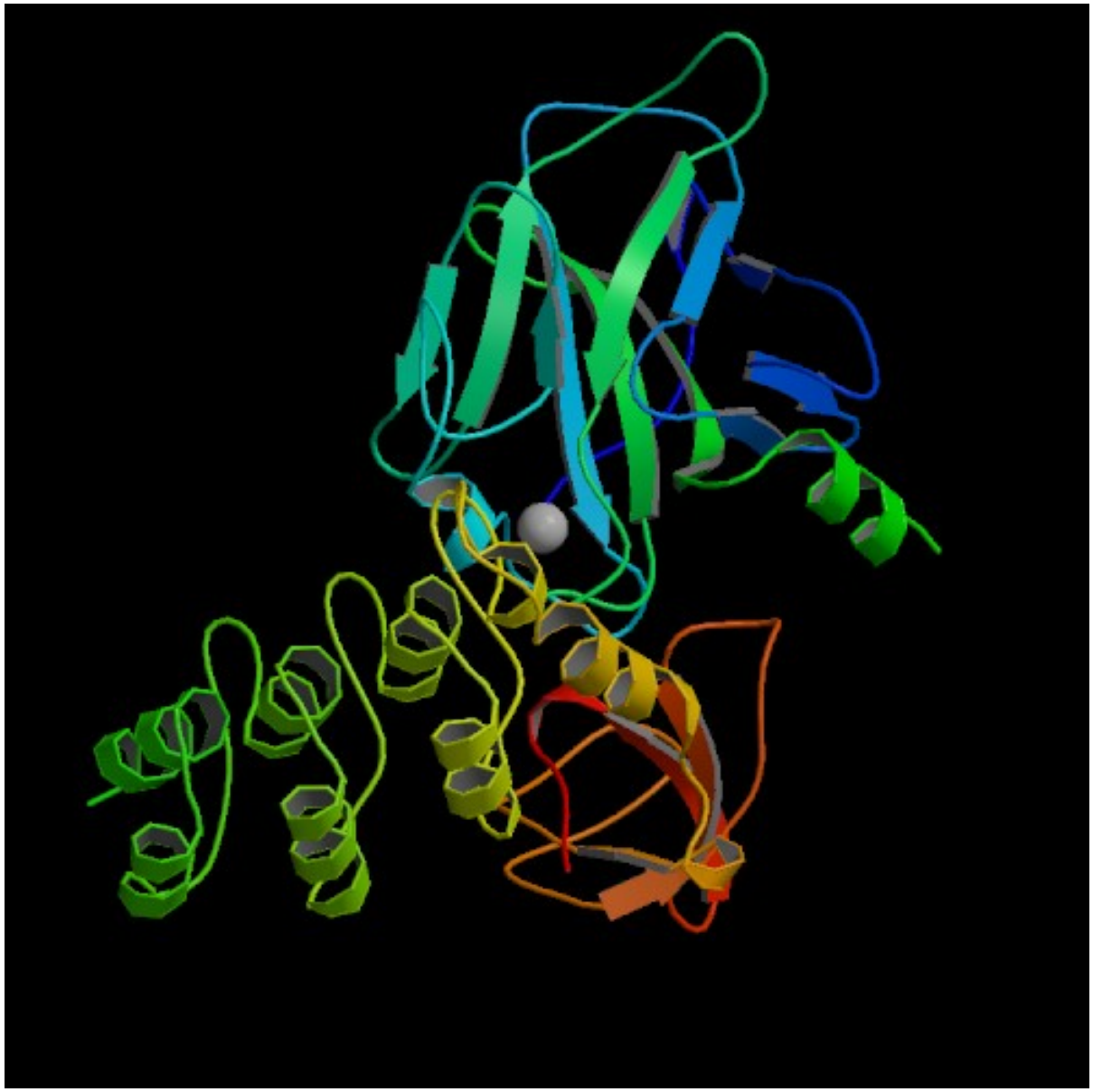


A ball-and-stick model of HIV II protease active site residues complexed with L-735,524 which is an orally bioavailable inhibitor of the HIV protease

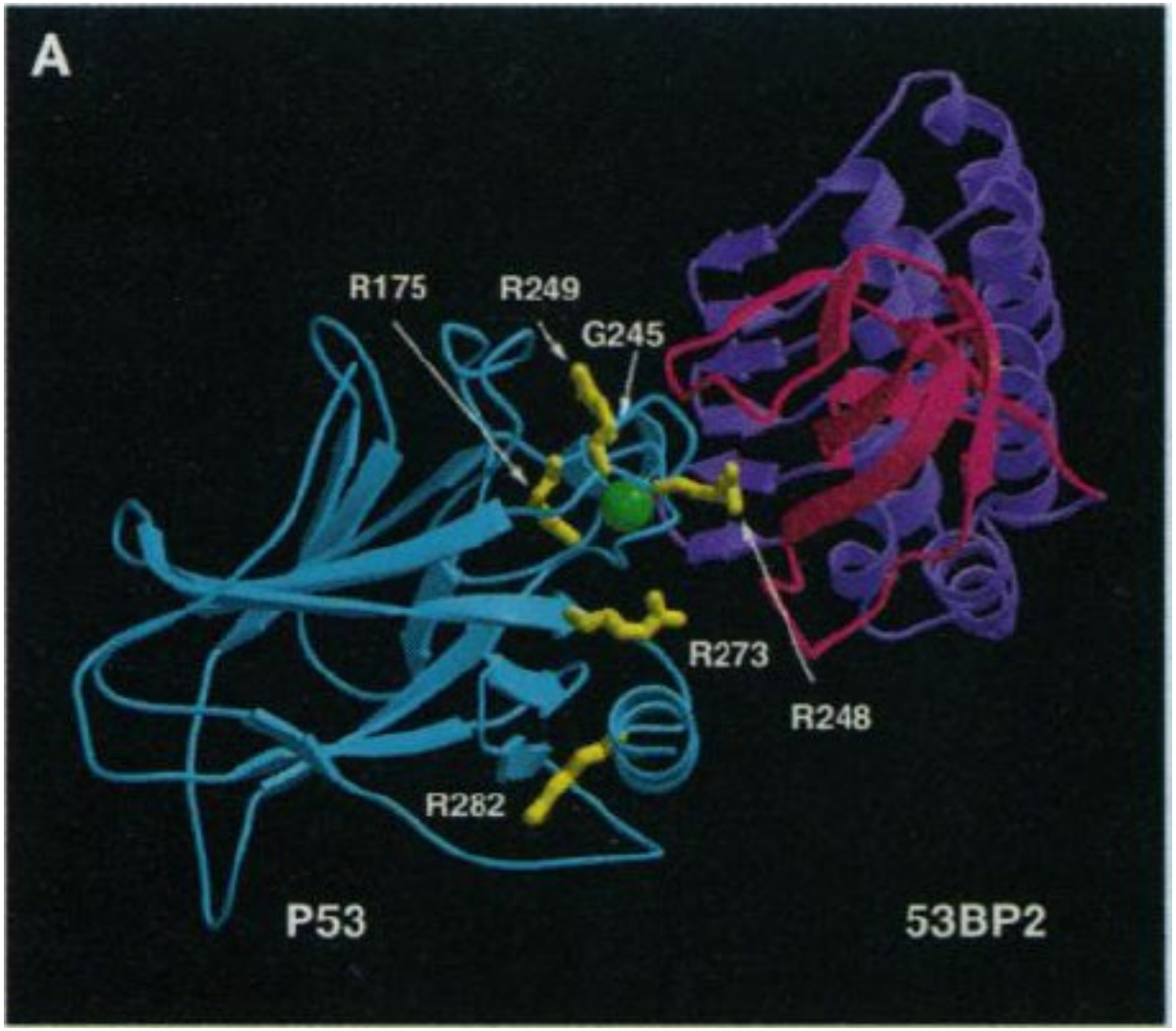
J.Biol.Chem. v269 pp.26344-26348 , 1994

Biotechnology in 3-Dimensions

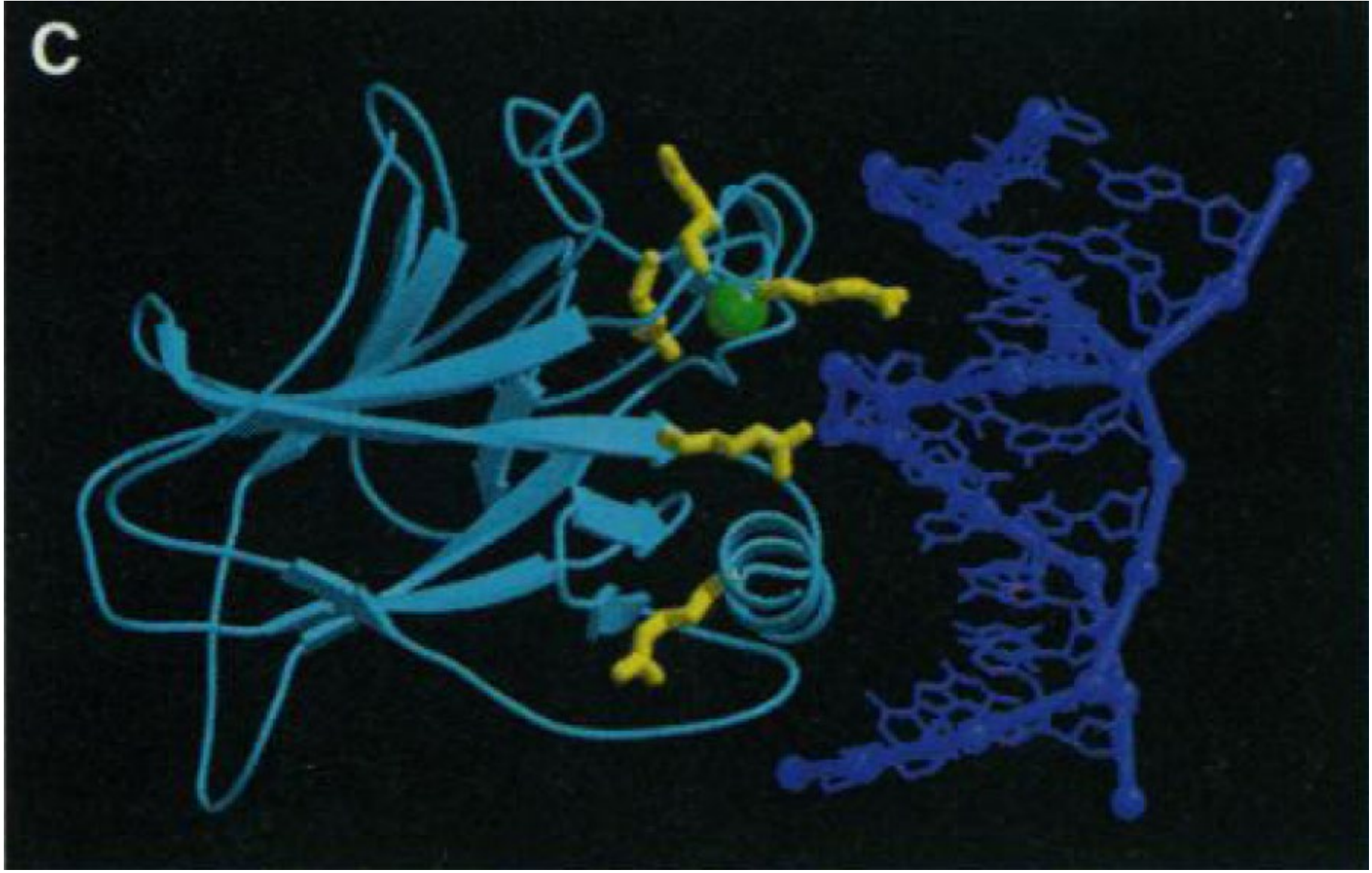
- Useful for identifying inhibitors & agonists
- Useful for identifying protein-protein and/or protein/DNA interactions



From Protein Data Bank (PDB) file **1YCS**
Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2.
Science v274 pp.1001-1005 , 1996



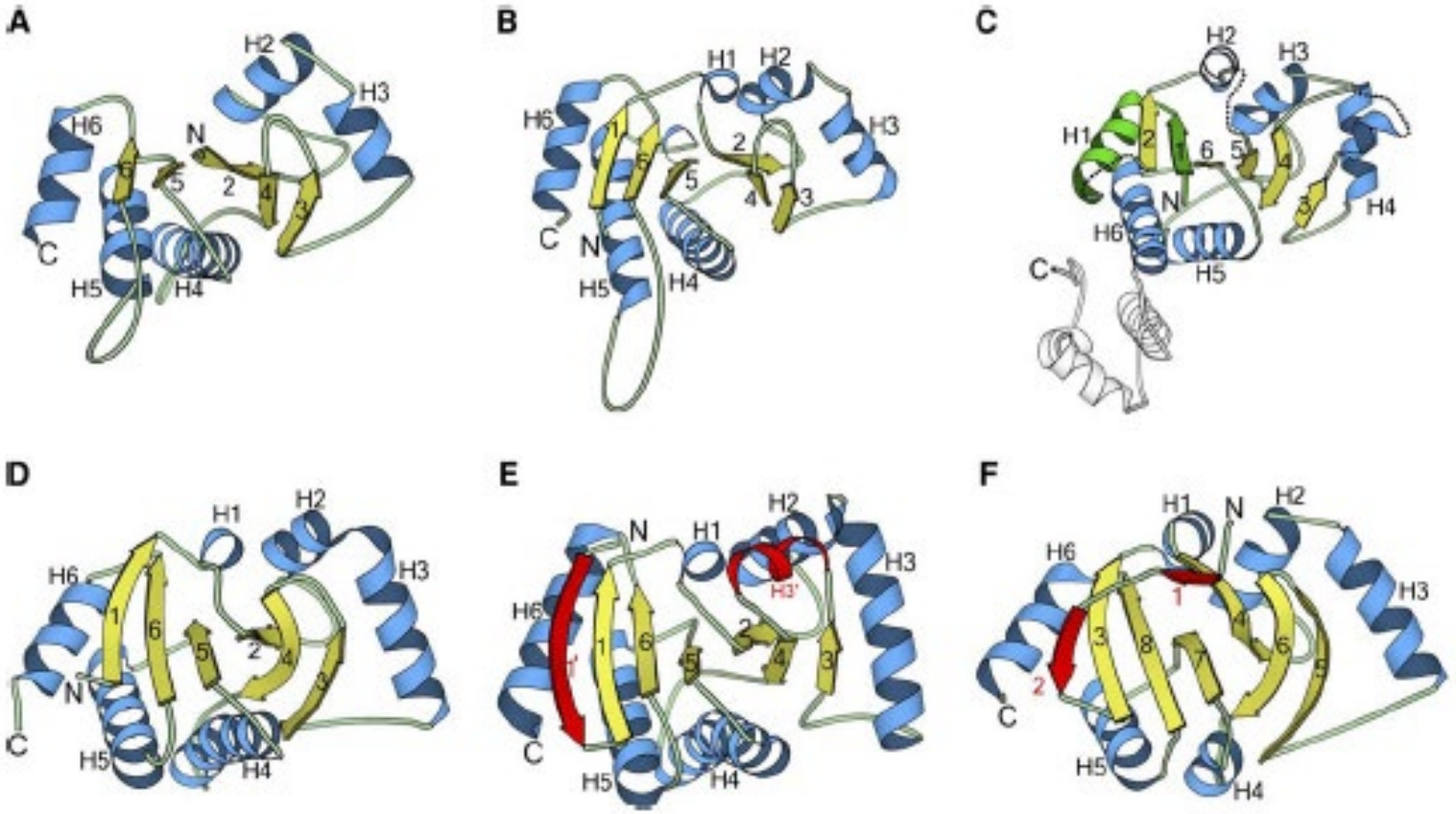
From Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2.
Science v274 pp.1001-1005 , 1996



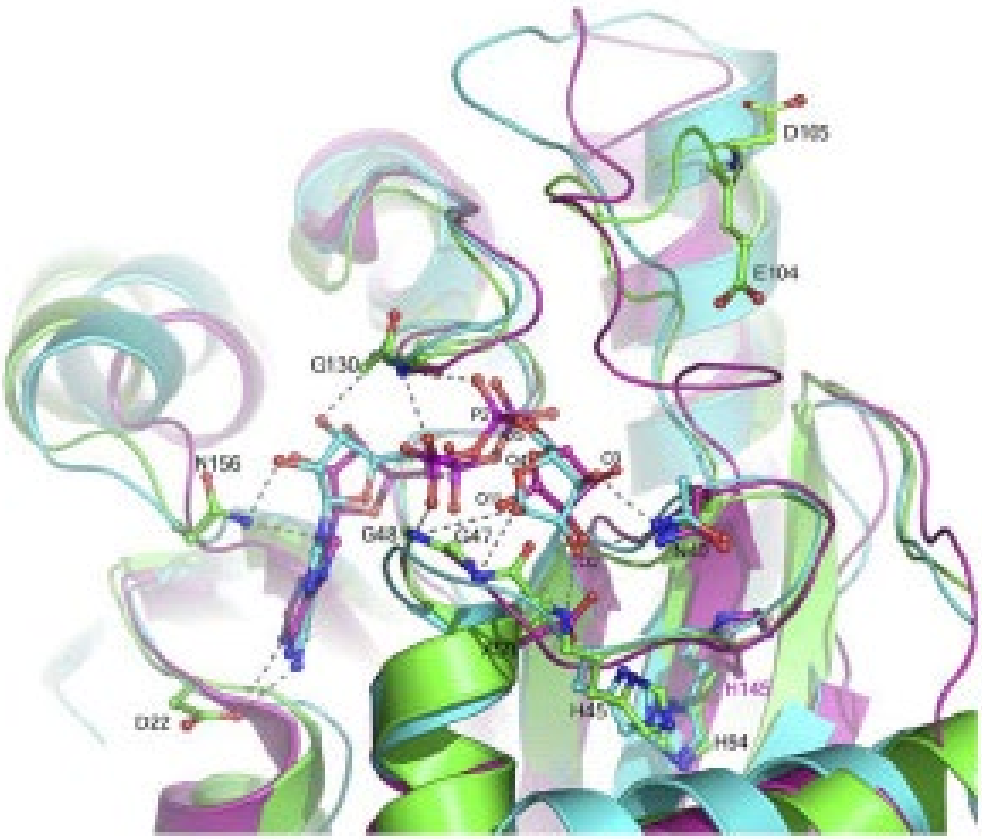
From Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2.
Science v274 pp.1001-1005 , 1996

Biotechnology in 3-Dimensions

- Useful for identifying inhibitors & agonists
- Useful for identifying protein-protein interactions
- Useful in proteomics



Identifying a fold of the SARS ADRP Domain.
A is a bovine Leu-aminopeptidase, B is *E. coli* pepA, C is a yeast Appr phosphatase, D is *E. coli* hypothetical protein Er58, E is *Archeoglobus fuldiges* AF1521, and F is ADRP domain of SARS nsp3.
Structure v13, pp.1665-1675, 2005



Identifying the active site of the SARS ADRP Domain.
yeast Appr phosphatase homolog is in purple, *Archeoglobus fuldiges* AF1521 in cyan, and SARS nsp3 in green with a ball-and-stick depiction of ADP-ribose in the active site.
Structure v13, pp.1665-1675, 2005

Biotechnology in 3-Dimensions

- Useful for identifying inhibitors & agonists
- Useful for identifying protein-protein and/or protein/DNA interactions
- Useful in proteomics
- Useful in producing “designer” proteins

Trilateral Co-operation
Biotechnology Project WM4
Protein 3-Dimensional (3d) structure related claims

- November, 2002
- EPO, JPO, and USPTO input
- Addressing “increasing numbers of applications claiming inventions related to … 3-D structural information”

35 USC § 101
Statutory Classes of Invention

35 USC § 101 reads

"Whoever invents or discovers any new and useful **process**, **machine**, **manufacture**, or **composition**, or any new and useful improvement thereof, may obtain a patent therefore..." (emphasis added)

Categories =process
machine
manufacture
composition of matter

35 USC § 101
Statutory Classes of Invention - Examples

Claim 1. A data array comprising the atomic coordinates of protein P as set forth in Figure 1.

The 3-D coordinates of a protein constitute **nonfunctional descriptive material**.

Claim 2. A computer model of protein P generated from the data array of Claim 1.

Claim 3. A computer-readable storage medium encoded with the data array of Claim 1.

Claim 4. A computer comprising the data array of Claim 1 stored in memory.

Claim 5. The computer of Claim 4, additionally comprising executable code for:
(a) displaying the data array as a 3-dimensional model;
(b) analyzing the binding site of the model of protein P;
(c) screening *in silico* a library for small molecules that fit into said binding site; and
(d) controlling a unit for assaying the small molecules determined in step (c) in a protein P binding assay.

Claims 1-3 paraphrased from Trilateral Project WM4, Cases 1 and 2. See also MPEP 2106

35 USC § 101
Statutory Classes of Invention - Examples

Claim 6. A data array comprising the atomic coordinates of protein P as set forth in Figure 1.

Claim 7. An isolated protein P having the structure defined by the structural coordinates of the data array of Claim 6.

Claim 8. A pharmacophore having a spatial arrangement of atoms defined by the binding pocket identified in the data array of Claim 6.

A **pharmacophore** is a description of a generalized concept of molecular features in terms of information on spatial arrangement of chemical elements (e.g. hydrophobic groups, ionizable groups, H bond donors/acceptors, etc.)

Claim 9. An isolated compound or its salt defined by the pharmacophore of Claim 8.

Claims paraphrased from Trilateral Project WM4, Cases 1, 3, and 8

35 USC § 112, first paragraph
Written Description

35 USC § 112, first paragraph reads

"The specification shall contain a **written description of the invention**, and of the manner and process of making and using it, in such full, clear, concise, and exact terms **as to enable any person skilled in the art** to which it pertains, or with which it is most nearly connected, **to make and use the same** and shall set forth the best mode contemplated by the inventor of carrying out his invention."(emphasis added)

35 USC § 112, first paragraph
Written Description - Examples

Claim 10. An isolated and purified protein P having the structure defined by the structural coordinates as shown in Figure 1.

Figure 1 teaches a "complete" 3D structure of full-length protein P.

Claim 11. An isolated and purified protein P having:

a) a molecular weight of 315 kD as measured by SDS-PAGE,

b) a pI of 7.5,

c) an N-terminal amino acid sequence of SEQ ID NO:10, and

d) the activity of full length protein P.

Claim 10 paraphrased from Trilateral Project WM4, Case 3

35 USC § 112, first paragraph
Written Description - Examples

Claim 12. An isolated and purified protein P having at least the structure defined by the binding pocket amino acids identified in Figure 2.

Figure 2 teaches a *partial* 3D structure of protein P, limited to the binding pocket amino acids, which are about only 10% of the entire protein.

Claim 13. An isolated and purified protein P having:

a) a protease fragment with a molecular weight of 31 kD as measured by SDS-PAGE,

b) a protease fragment with a pI of 7.5,

c) an N-terminal amino acid sequence of SEQ ID NO:10, and

d) the activity of full length protein P.

Claim 12 paraphrased from Trilateral Project WM4, Case 5

35 USC § 112, first paragraph
Enablement

35 USC § 112, first paragraph reads

"The specification shall contain a written description of the invention, and of **the manner and process of making and using it**, in such full, clear, concise, and exact terms **as to enable any person skilled in the art** to which it pertains, or with which it is most nearly connected, **to make and use the same** and shall set forth the best mode contemplated by the inventor of carrying out his invention."

(emphasis added)

35 USC § 112, first paragraph
Enablement - Examples

Claim 14. The crystalline form of protein P having unit cell dimensions a=4.0 nm, b=7.8nm, and c=11.0nm.

The specification teaches the recombinant expression and purification of the claimed protein P as defined by **SEQ ID NO:2** (which includes a His-tag for ease of purification in *E. coli*). This purified protein sample was then subjected to *clearly described* crystallization conditions to produce x-ray quality crystals of a particular **unit cell dimension** (e.g. the size of the repeating unit in the ordered array) and **space group** $P2_12_1$ (e.g. the organization of the repeating unit).

Claim paraphrased from Trilateral Project WM4, Case 4

35 USC § 112, first paragraph
Enablement - Examples

Claim 15. An isolated and purified protein having the sequence shown in SEQ ID NO:1.

Claim 16. The protein of Claim 15 in crystalline form.

Claim 17. The protein of Claim 15 in soluble form.

Claims paraphrased from Trilateral Project WM4, Cases 3 and 4

35 USC § 112, first paragraph
Enablement - Examples

Claim 18. A method of identifying compounds that bind protein P comprising:

(a) obtaining a 3-D molecular model of protein P as shown in Figure 1;

(b) reducing said model to a 3-D molecular model of the binding pocket of protein P as shown in Figure 2;

(c) comparing the model of (b) with a library of 3-D molecular models representing structures of candidate compounds to electronically screen said library;

(d) identifying candidate compounds whose structures electronically fit in the model of (b) as compounds that can bind protein P; and

(e) assaying the binding of candidate compounds identified in step (d) using purified protein P;

to thereby identify compounds that bind protein P.

Claim paraphrased from Trilateral Project WM4, a combination of Cases 6 and 7

35 USC § 102 and 103
Prior Art

35 USC § 102(b) reads

"A person shall be entitled to a patent **unless** –

the invention was patented or **described** in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year **prior** to the date of application for patent in the United States." (emphasis added)

35 USC § 103(a) reads

"A **patent may not be obtained** though the invention is not identically disclosed or described as set forth in section 102 of this title, **if the differences** between the subject matter sought to be patented and the prior art are such that the subject matter as a whole **would have been obvious at the time the invention** was made to a person having ordinary skill in the art to which said subject matter pertains."

(emphasis added)

35 USC § 102 and 103
Prior Art – Examples

Claim 19. An isolated and purified protein P having the structure defined by the structural coordinates as shown in Figure 1.

Claim 20. The protein of Claim 19 in crystalline form.

Claims paraphrased from Trilateral Project WM4, Case 3 and 4

35 USC § 102 and 103
Prior Art – Examples

Claim 21. A method of identifying compounds that bind protein P comprising:

(a) obtaining a model of protein P as shown in Figure 1;

(b) using said model in a method of rational drug design to identify candidate compounds that can bind protein P; and

(c) assaying the binding of candidate compounds identified in step (b) using purified protein P;

to thereby identify compounds that bind protein P.

Claim paraphrased from Trilateral Project WM4, a combination of Cases 6 and 7

References

- Trilateral Co-operation Biotechnology Project on 3-dimensional proteins http://www.trilateral.net/projects/biotechnology/protein_3d/

- USPTO Guidelines for Computer-related inventions http://www.uspto.gov/web/offices/pac/compexam/comguide.htm

- In re Gulack, 217 USPQ 401 (Fed. Cir. 1983)
- In re Ngai, 70 USPQ2d 1862 (Fed. Cir. 2004)
- In re Lowry, 32 USPQ2d 1031 (Fed. Cir. 1994)
- In re Warmerdam, 31 USPQ2d 1754 (Fed. Cir. 1994)
- State Street Bank & Trust Co. v. Signature Financial Group Inc., 47 USPQ2d 1596 (Fed. Cir. 1998)

- NCBI Structure Database http://www.ncbi.nlm.nih.gov/Structure/
- Protein Data Bank (PDB) http://www.rcsb.org/pdb/home/home.do

Acknowledgements

Nashaat Nashed
David Steadman
Suzanne Noakes
Alexander Kim
Jae Wan Lee
Ardin Marschel
Jean Witz