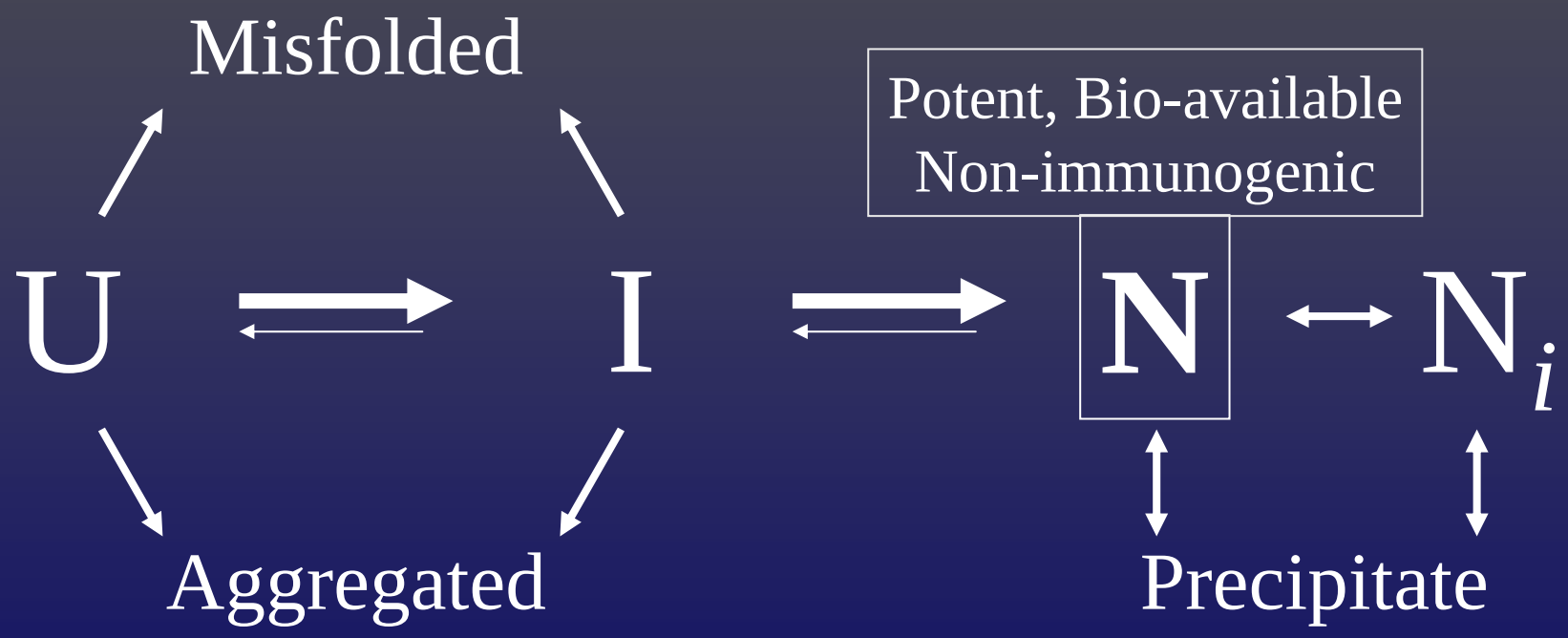
Optical Spectroscopy for Protein Secondary Structure Evaluation

Keith A. Oberg, Ph.D.
Medical Research Products A
Advanced Bionics
& MannKind Corporation

- Optical Spectroscopy can help demonstrate equivalence
- Optical Spectroscopy can help *achieve* equivalence

MEDICAL research PRODUCTS

Conformational States of Proteins



Each state can have a unique spectroscopic signature

A Context for Optical Spectroscopy

- Familiarity with the protein is critical
  - Initial handling & pilot studies
  - Formulation development
  - Process development
- Rapid & ‘simple’ tools are crucial
  - Enables progress in initial stages
  - Provides insight to the effect of environment
  - Speeds resolution of development issues
  - Improves in-process assay turnaround

MEDICAL research PRODUCTS

A Context for Optical Spectroscopy

- Rapid
- Repeatable
- Inexpensive
- Nondestructive
- In-situ analysis possible
- Part-11 compliance
- Validatable
- Valuable screening tool - reduces the need for complex / slow methods

- Low resolution
- Not always simple
- Not always reliable
- Troubleshooting not trivial
- Learning curve
- Limiting sample requirements
- Orthogonal data essential

MEDICAL research PRODUCTS

Overview of Optical Methods & Application in Formulation

UV Absorbance

- Parameters
  - Intensity
  - Band Shape
- Sensitivity
  - Aromatic Residues
  - Disulfides
  - *Amide Bond
- Uses
  - Quantitation
  - Tertiary structure
  - Denaturation -C_m, Melts -T_m
- Characteristics
  - Reliable, simple, ubiquitous
  - Limited structural information

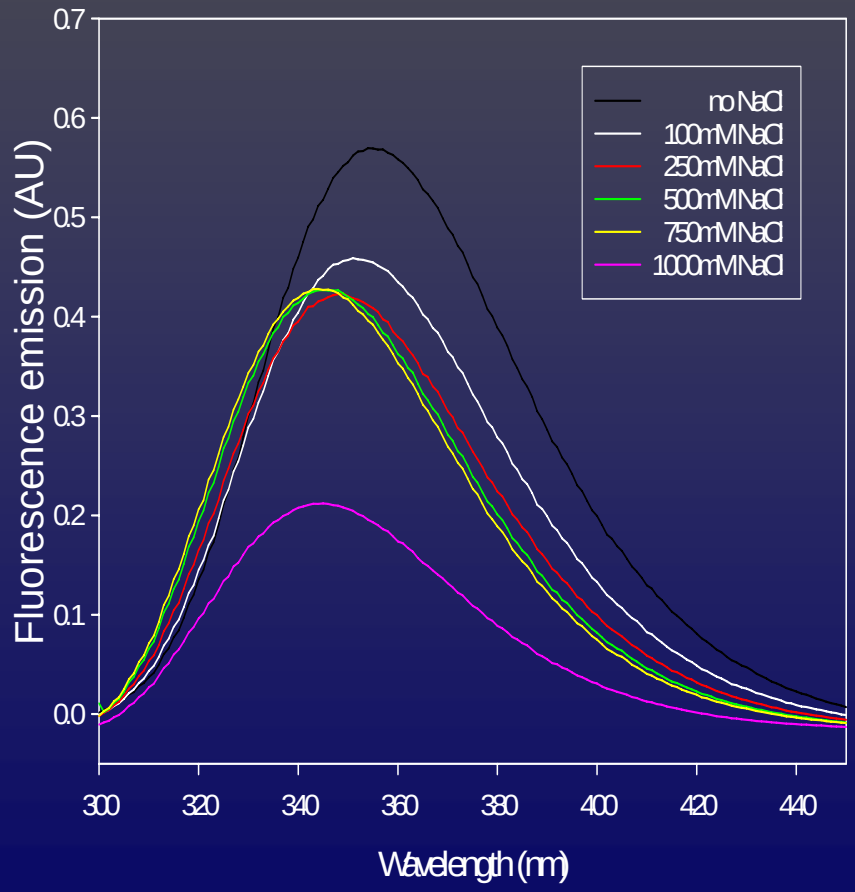
MEDICAL research PRODUCTS

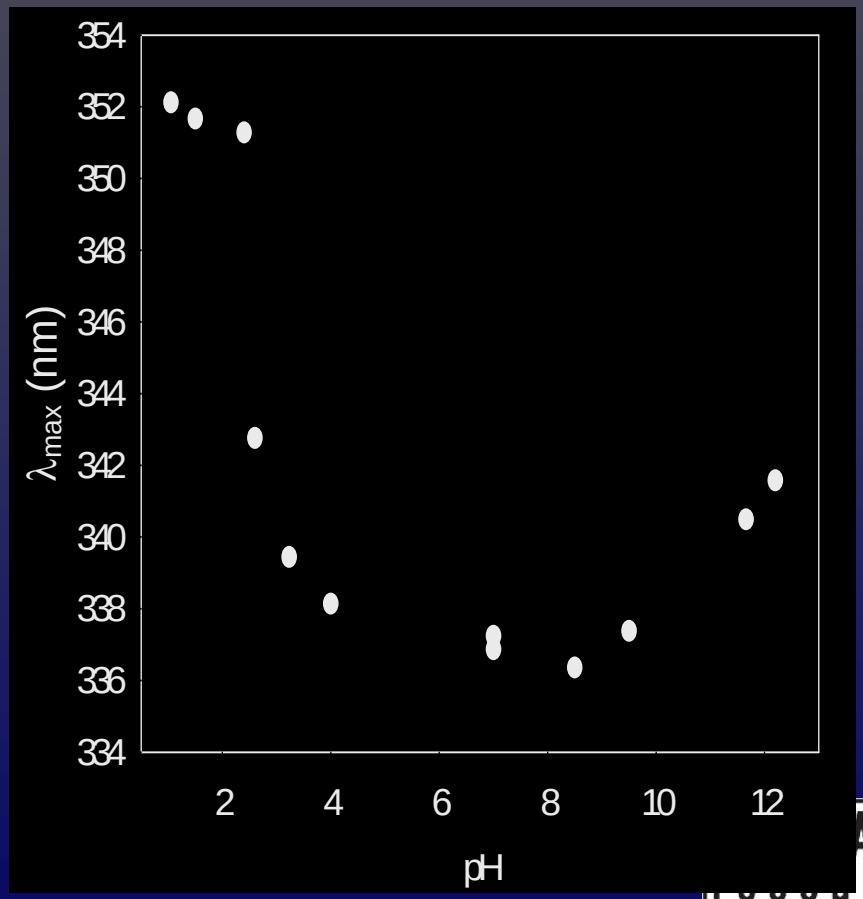
Fluorescence

- Parameters
  - Intensity
  - Wavelength
  - Anisotropy/polarization
- Sensitivity
  - Aromatic residues
  - Residue exposure
  - Residue proximity (FRET or quenching)
  - Cofactors
  - Dyes
- Uses
  - Denaturation -C_m, T_m
  - Tertiary or quaternary structure tracking
  - Hydrodynamic properties, residue mobility
  - Quenching
  - Accessible hydrophobics
- Characteristics
  - Protein-specific
  - Low concentration*

MEDICAL research PRODUCTS

Fluorescence Examples



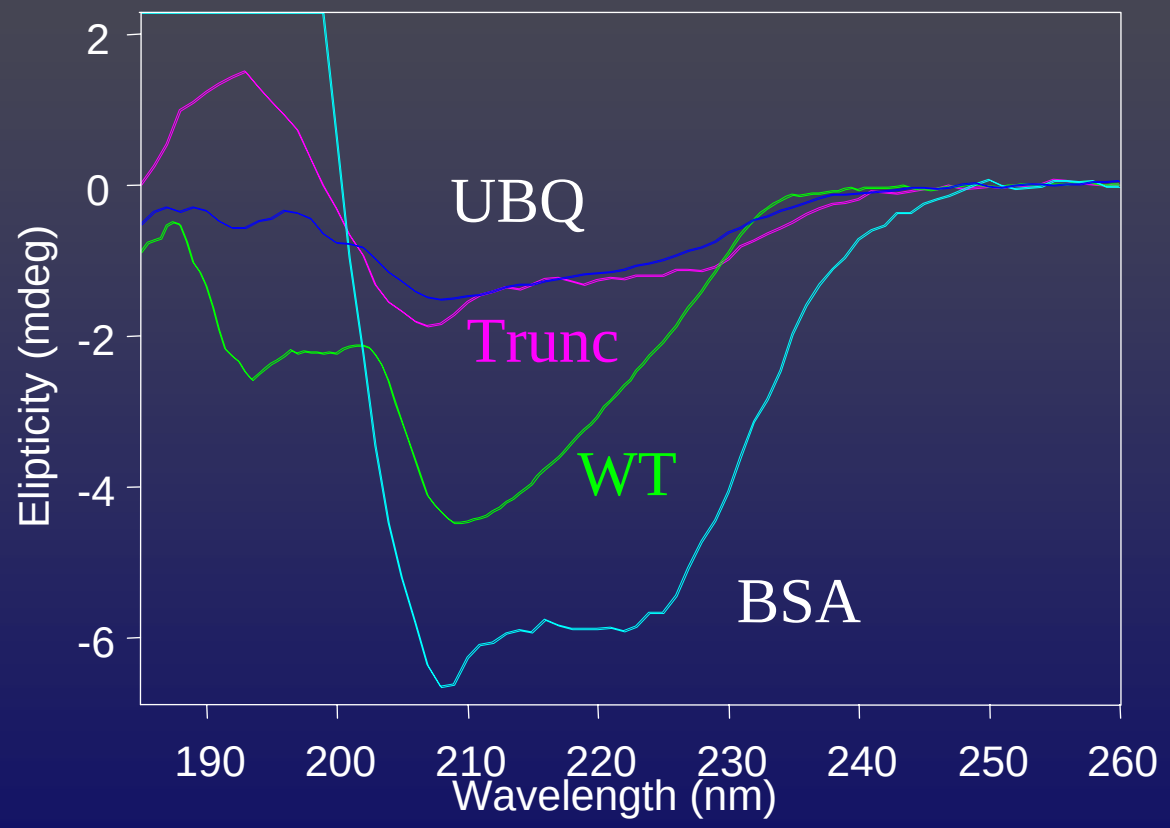


Circular Dichroism

- Parameters
  - Ellipticity
  - Band shape (wavelength)
- Sensitivity
  - 2° structure α, U
  - Chiral environment of aromatics & SS
- Uses
  - “Absolute” structure
  - structure comparisons
  - Cofactor binding & environment
- Characteristics
  - Small β-sheet signal
  - Newer equipment is better
  - Intensity critical
  - Reference set needed for *a priori* structure analysis
  - Far UV CD
    - Amide bands broad
    - Sensitive to UV absorbing solution components
    - Low concentration 0.01 - 5g/l, short path length better
    - aromatic contributions are protein specific
  - Near UV CD
    - >1g/l, longer PL better
    - protein specific

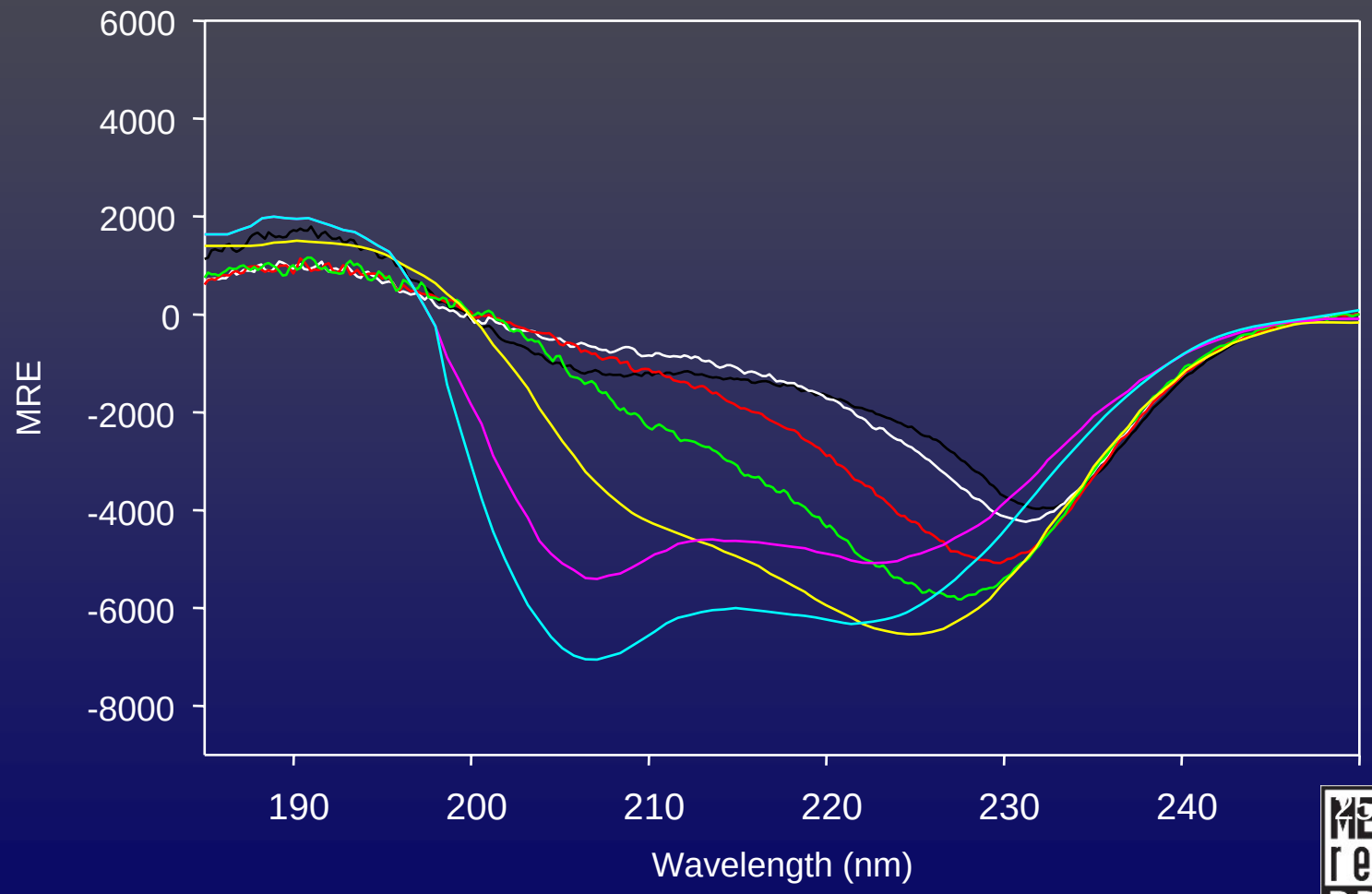
MEDICAL research PRODUCTS

CD Example 1



Truncated and full-length protein had different CD signals

CD Example 2



MEDICAL research PRODUCTS

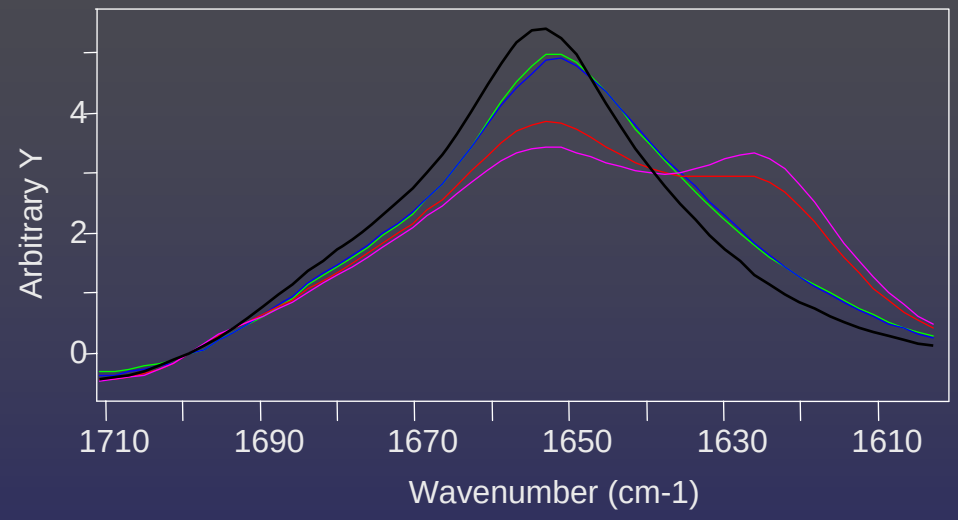
FTIR & (VCD)

- **Parameters**
  - Band-shape
  - Frequency (wavelength)
- **Sensitivity**
  - 2° Structure - β sheet, & helix/non-periodic
  - Side-chains
- **Uses**
  - Structure estimation
  - Secondary structural comparisons
  - Side-chain environment & conformation
  - Ligand binding
  - Conformational flexibility

- **Characteristics**
  - High protein concentrations (>5g/l)
  - Interference from carboxylate in amide I
  - Side chains and backbone signals overlap
  - Band *shape* critical
  - *A priori* structure estimation possible
  - Difference spectroscopy common & simple
  - Alternative solvents (D_2O) common
  - Transparency not required
  - Huge range of techniques accessible

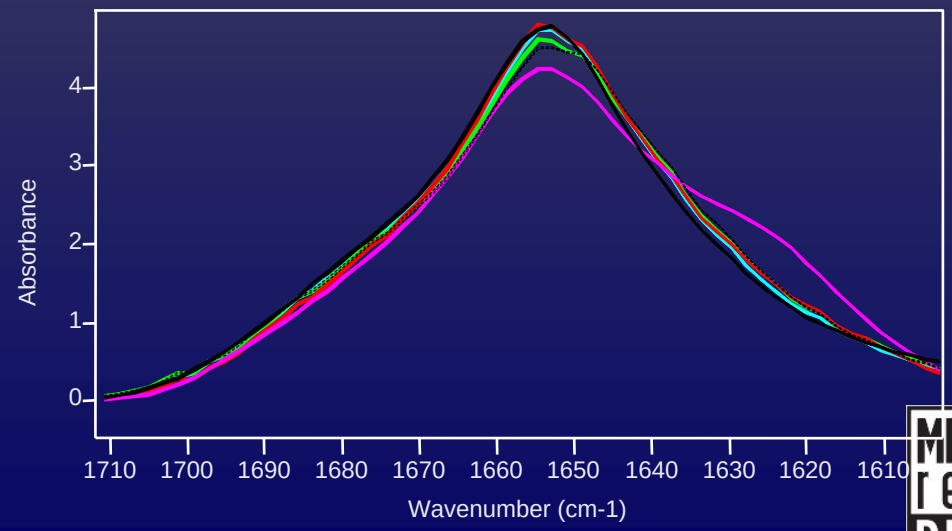
FTIR Example

Comparison of formulation pH



pH 7 black
pH 6 blue
pH 5 green
pH 3 red
pH 2 violet

pH Jumps simulate injection



MEDICAL research PRODUCTS

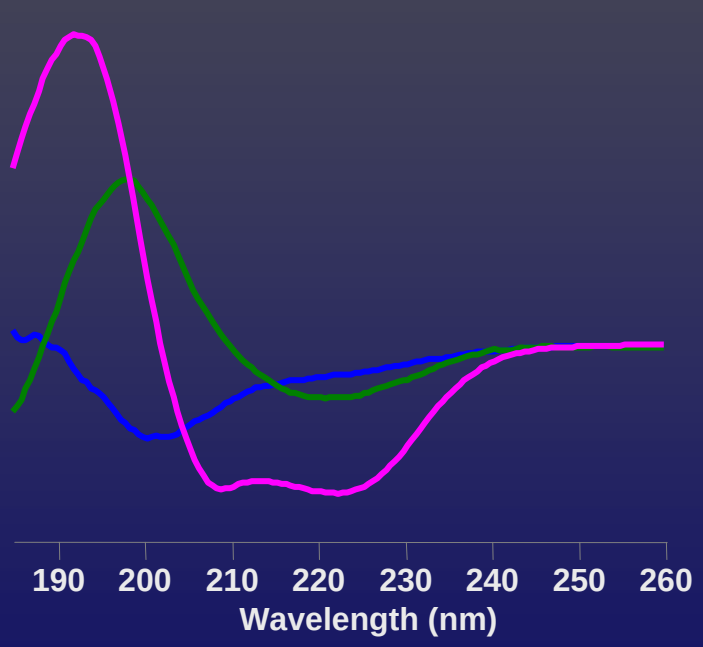
Minimal Optical Instrumentation for Protein Work

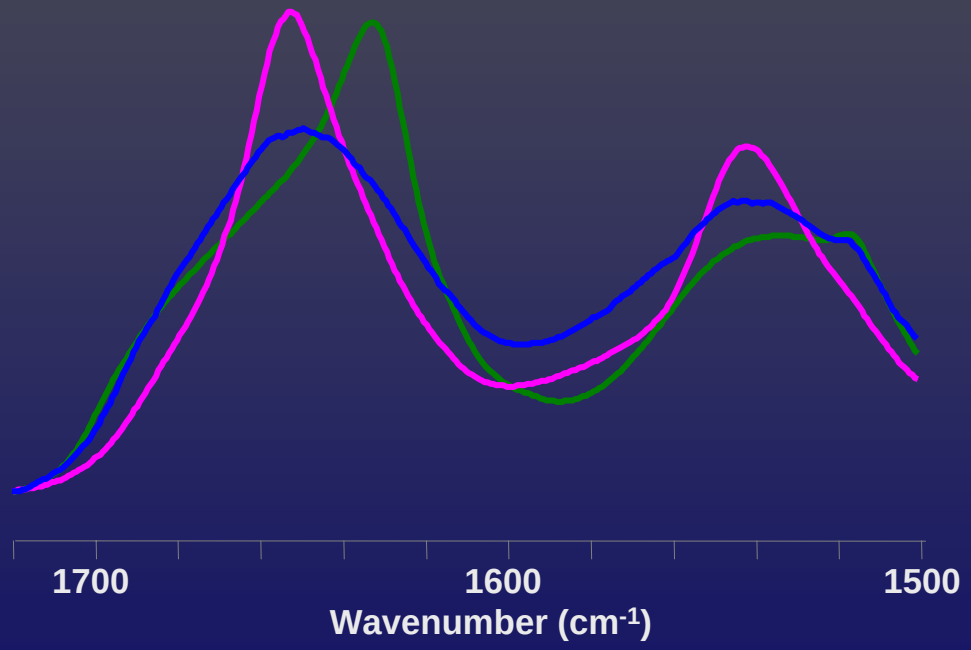
- CD-UV-Fluorescence
  - Rectangular cells, 10um – 1cm
  - Triangular cells – suspensions
  - Thermal control (T_m)
  - Titrator (C_m)
  - Integrated pH measurement
- FTIR
  - ATR –Ge IRE flow-cell (Horizontal)
  - Transmission – demountable CaF_2, 5um PTFE spacer

MEDICAL research PRODUCTS

Secondary Structure Determination With Optical Spectroscopy

Secondary Structure from Spectra



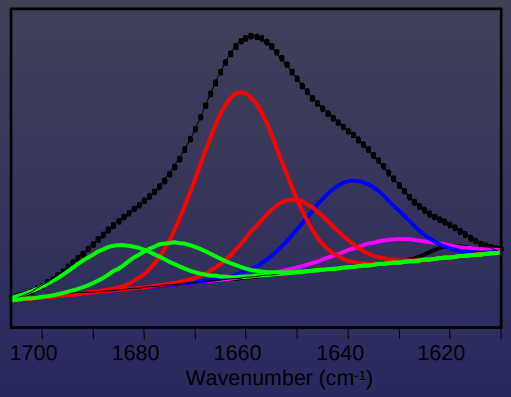


MEDICAL research PRODUCTS

2° Structure Determination from Optical Spectra

Curve Fitting:

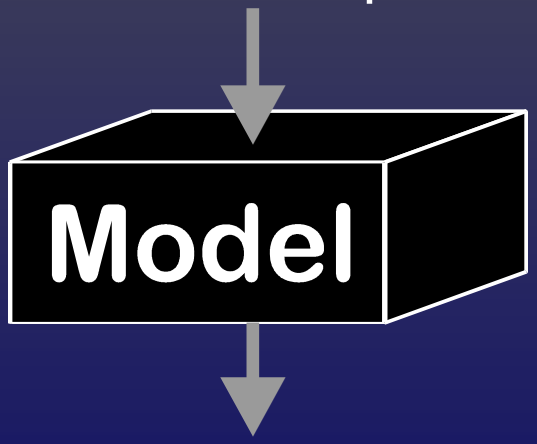
- Unknown spectrum



- Assign bands*
- Integrate areas
- Results (IR only)

Principal Components:

- Build basis set*
- Train algorithm
- Unknown spectrum

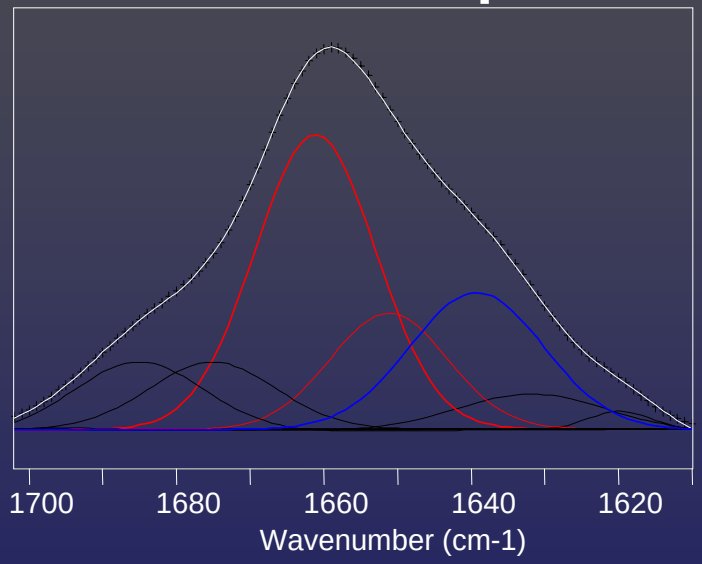


- Results (IR or CD)

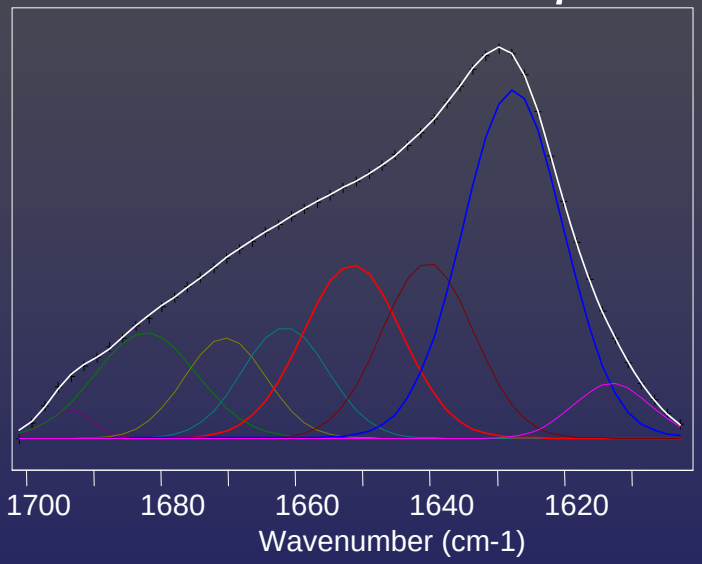
MEDICAL research PRODUCTS

Curve-Fitting of FTIR Spectra

Bacteriorhodopsin



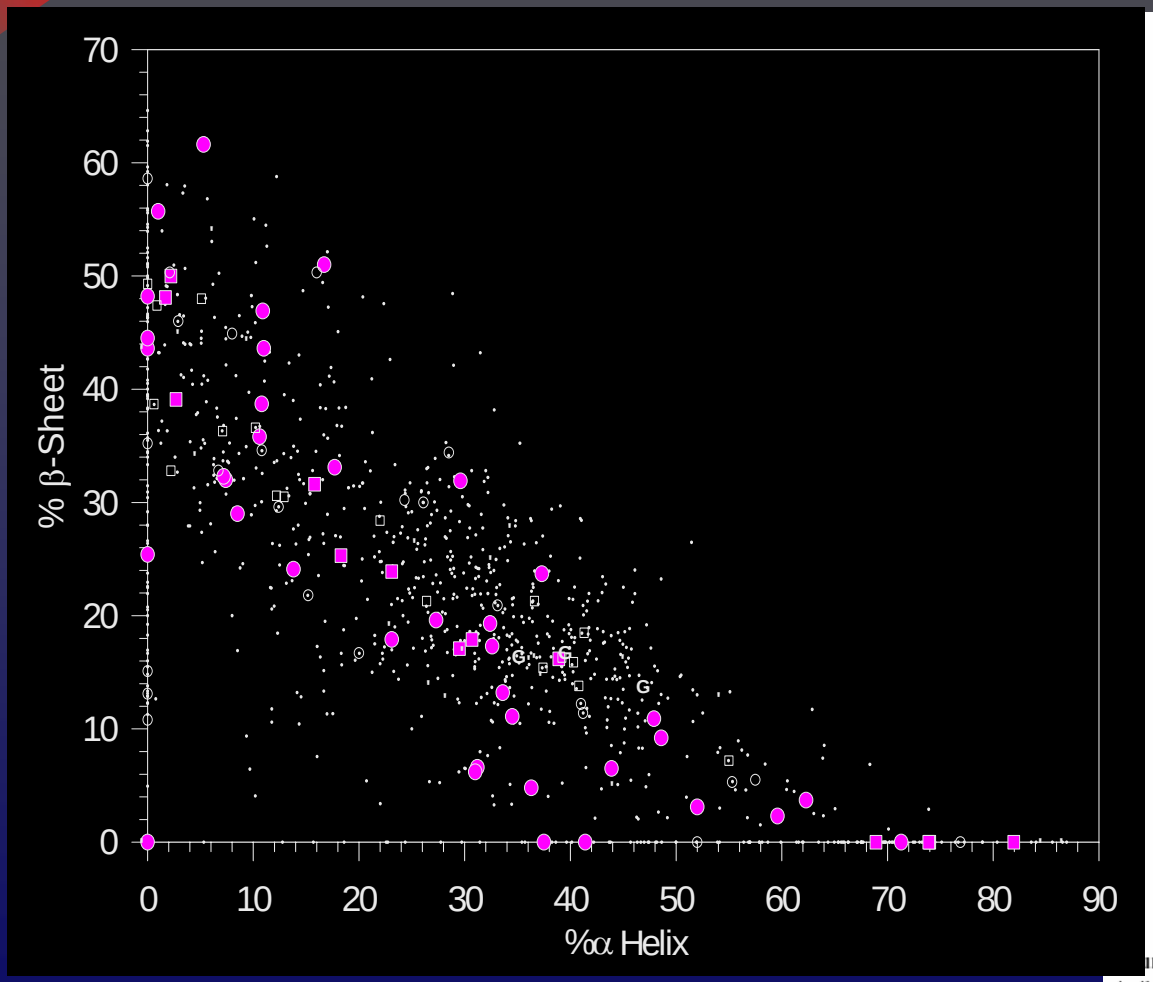
Interleukin 1β

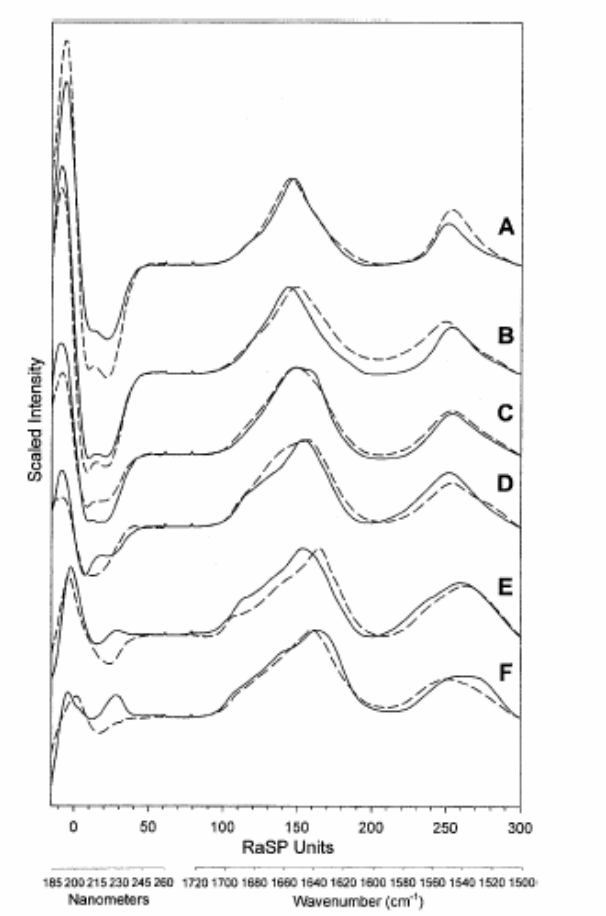


Method	% α-helix	% β-sheet	% turn
Curve fitting (FSD, auto) (Cabiaux et al., 1989)	51	18	21
Curve fitting (FSD, DR2) (Oberg and Fink, 1995)	52	20	24
Factor Analysis / RMR (Baumruk et al., 1996)	58	1	19
Electron Diffraction	**77**	**0**	**2**

Method	α-helix	β-sheet	turn	long loops
Curve fitting (FSD, DR2) Oberg et al. 1994	(25.4)	51.9	22.8	(25.4)
Curve Fitting Wilder et al. 1992	(20.6)	59.2	19.7	(20.6)
crystal structure	**0**	**56**	**20**	**24**

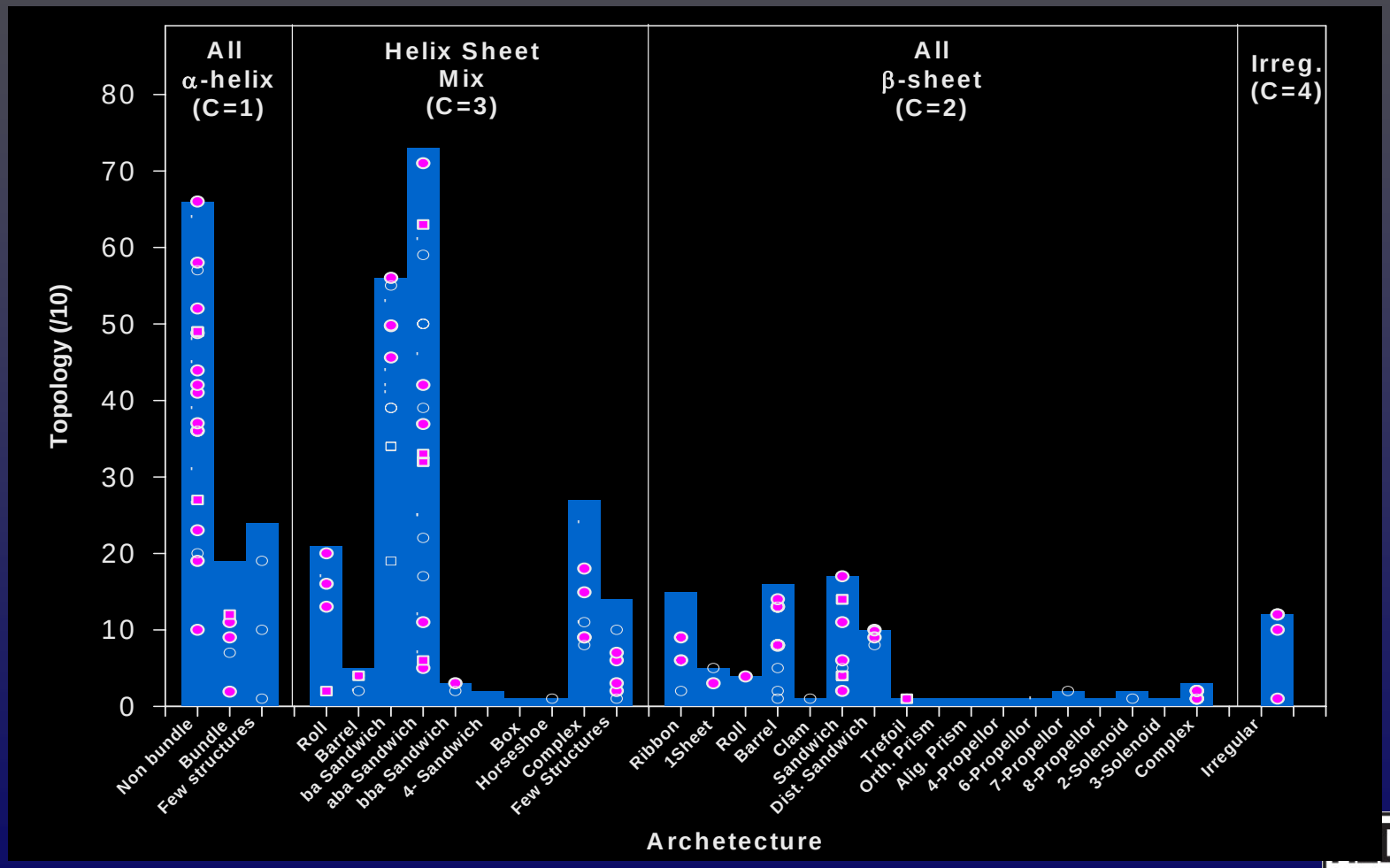
Properties of the RaSP50 set: HE Space & Secondary Structure





ure 5. Hybrid IR+CD spectra of some RaSP50 set protein pairs with similar secondary structure FCs but different folds. These demonstrate the spectral consequences of various differences in tertiary structure (Fig. 3).

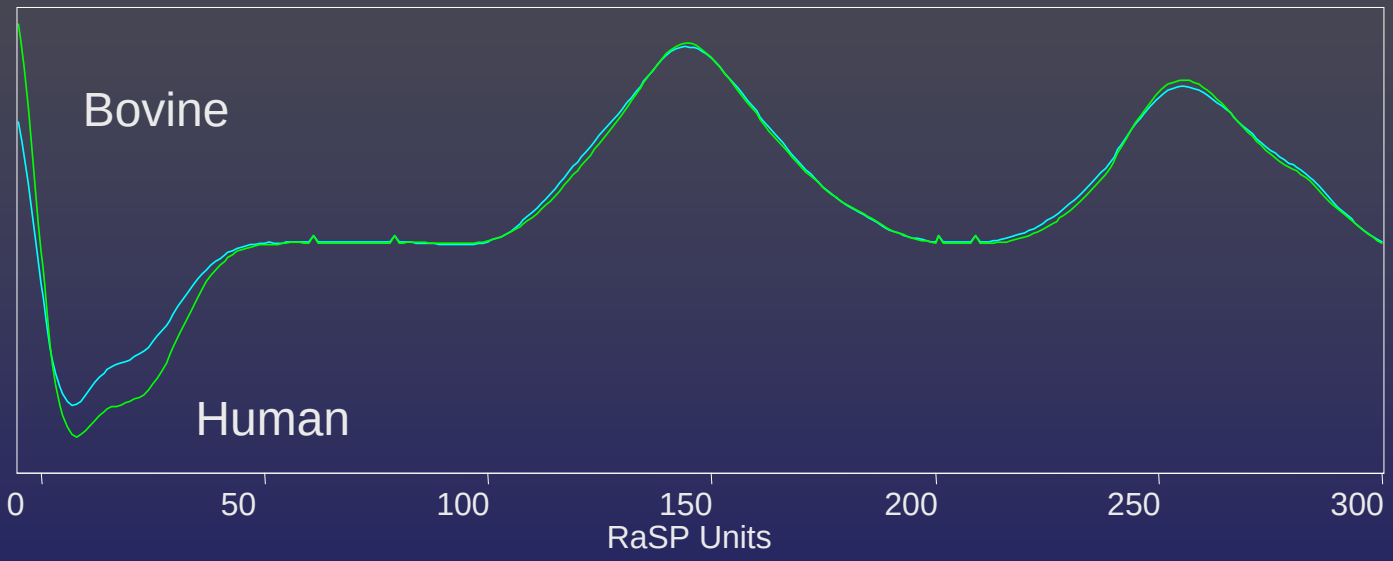
Properties of the RaSP50 set: CATH Space & Tertiary Structure



Example: Bacteriorhodopsin

Method	% α-helix	% β–sheet	% turn
crystal structure	77	0	2
FSD and **curve fitting** (Cabiaux et al., 1989)	51	18	21
FSD, deriv. & **curve fitting** (Oberg and Fink, 1995)	52	21	19
Factor Analysis / RMR (Baumruk et al., 1996)	58	1	18
Partial Least Squares TF-ATR, 32 proteins	67	0	10
RaSP (PLS, TF-ATR)	75	0	11

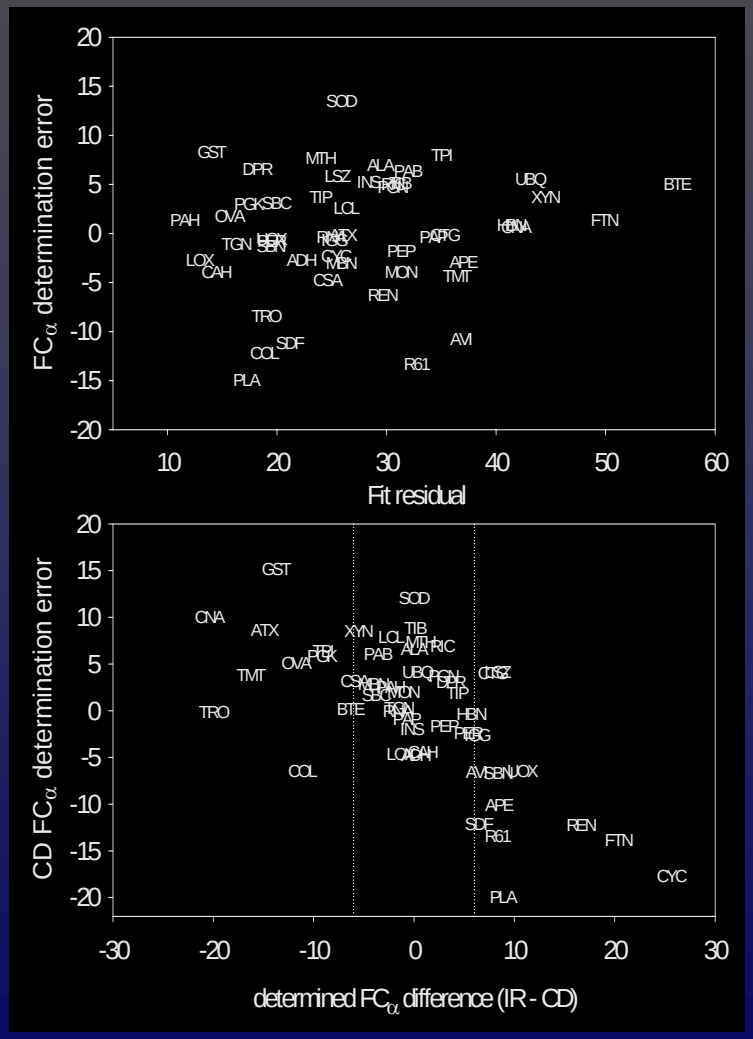
The use of Hybrid IR-CD Spectra



		Human α-lactalbumin (in RaSP)				Bovine α-lactalbumin (Not in RaSP)			
State	Data	α-Helix	β-Sheet	Turn	Other	α-Helix	β-Sheet	Turn	Other
Native	IR	35.9	2.9	15.1	42.9	31.0	13.4	14.3	40.3
Native	CD	25.3	11.5	12.6	39.8	26.1	14.2	12.9	39.1
Native	IR+CD	29.0	4.0	17.5	46.6	27.6	11.9	16.0	43.2
Native	Actual	**30.1**	**6.5**	**18.7**	**44.7**	**30.6**	**6.5**	**15.9**	**47.1**
A	IR	48.9	1.6	15.7	33.7	39.0	4.2	15.9	40.0
A	CD	42.2	7.4	12.3	37.5	29.3	14.1	12.9	42.5
A	IR+CD	43.8	0.0	15.7	41.9	31.9	1.6	17.0	48.0

MEDICAL research PRODUCTS

The Performance of the Rasp50 Set in Secondary Structure Analysis

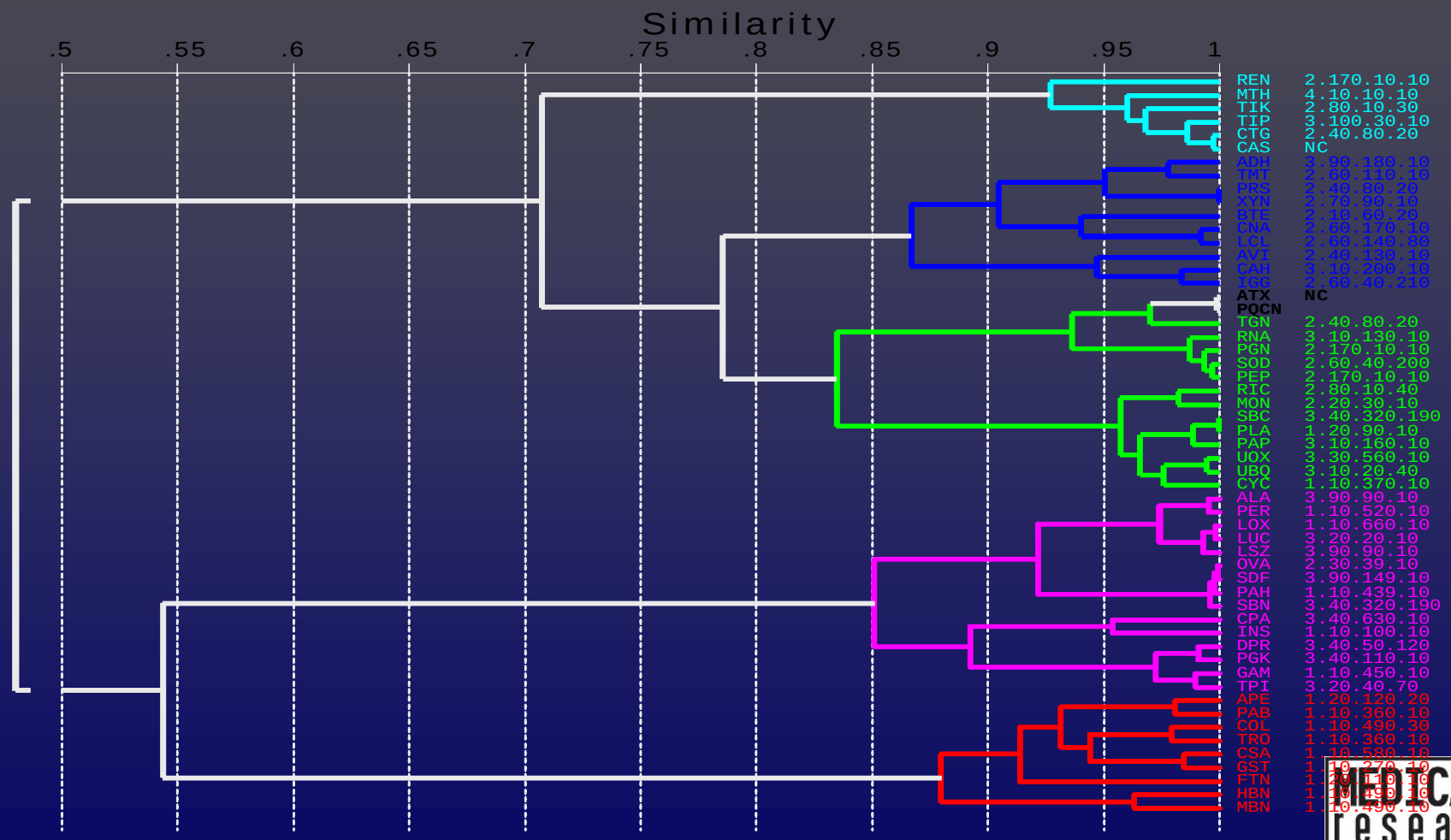


Data	α-helix RMS error	α-helix ζ_{subset}	β-sheet RMS error	β-sheet ζ_{subset}
	$\lvert\Delta IRCD_{det}\rvert < 6\%$			
IR alone	4.82	4.24	7.19	2.65
CD alone	4.46	4.58	7.62	2.50
IR+CD	4.46	4.59	6.26	3.04
	$\lvert\Delta IRCD_{det}\rvert > 6\%$			
IR alone	8.59	2.78	6.90	2.44
CD alone	9.60	2.49	10.26	1.64
IR+CD	6.54	3.66	6.94	2.43
max(CD/IR)	6.77	3.53	6.76	2.50

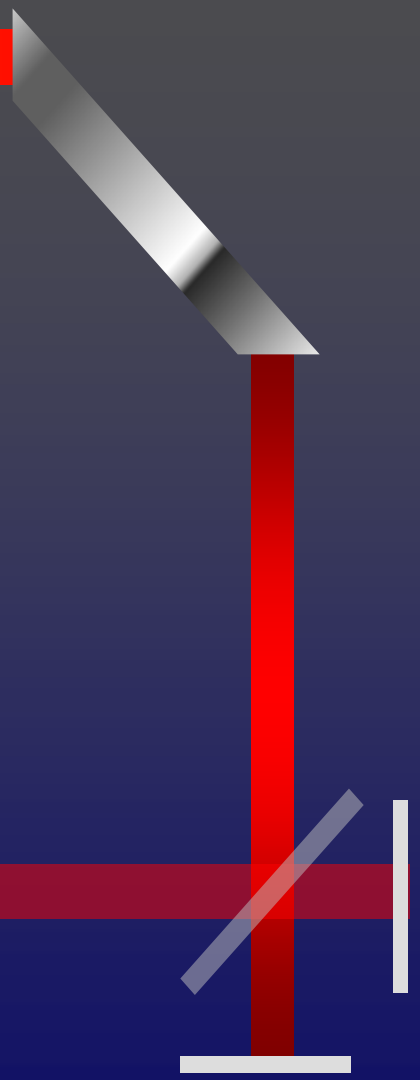
- RaSP analysis accuracy can be evaluated

MEDICAL research PRODUCTS

Cluster Analysis



Optical Spectroscopy for Determining Secondary Structural Equivalence



Curve-Fitting of FTIR Spectra

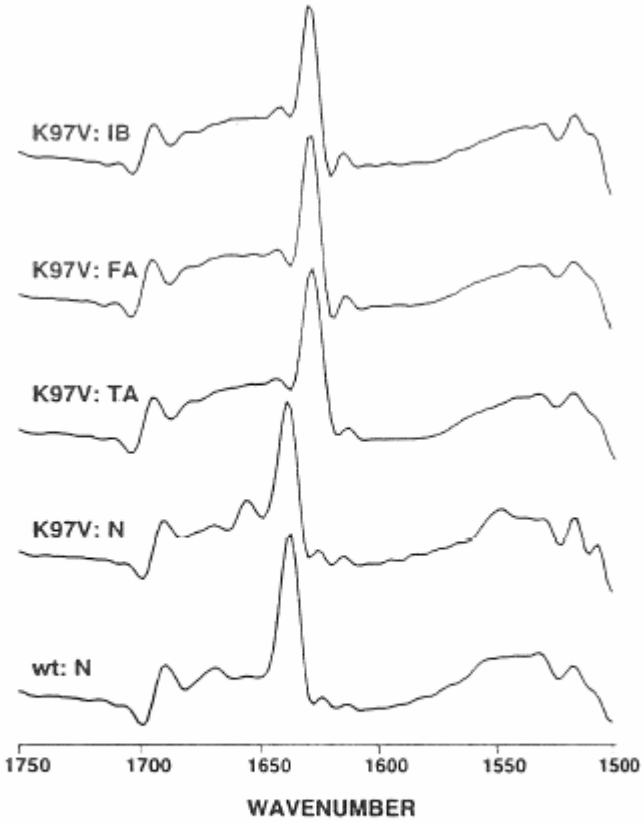


FIGURE 2: Comparison of the interleukin-1β amide I (1600–1700 cm^{-1}) and II (1500–1600 cm^{-1}) bands from ATR-FTIR after Fourier self-deconvolution (FSD) resolution enhancement. IB is the inclusion body, FA is folding aggregate, TA is thermal aggregate, K97V:N is native K97V, and wt:N is native wild type. The major peak (1625 or 1637 cm^{-1}) corresponds to β-sheet or extended structure.

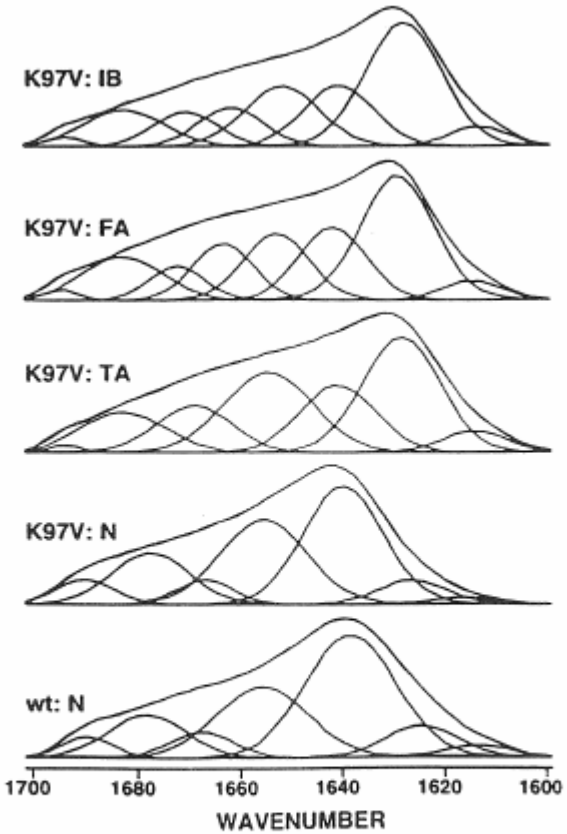


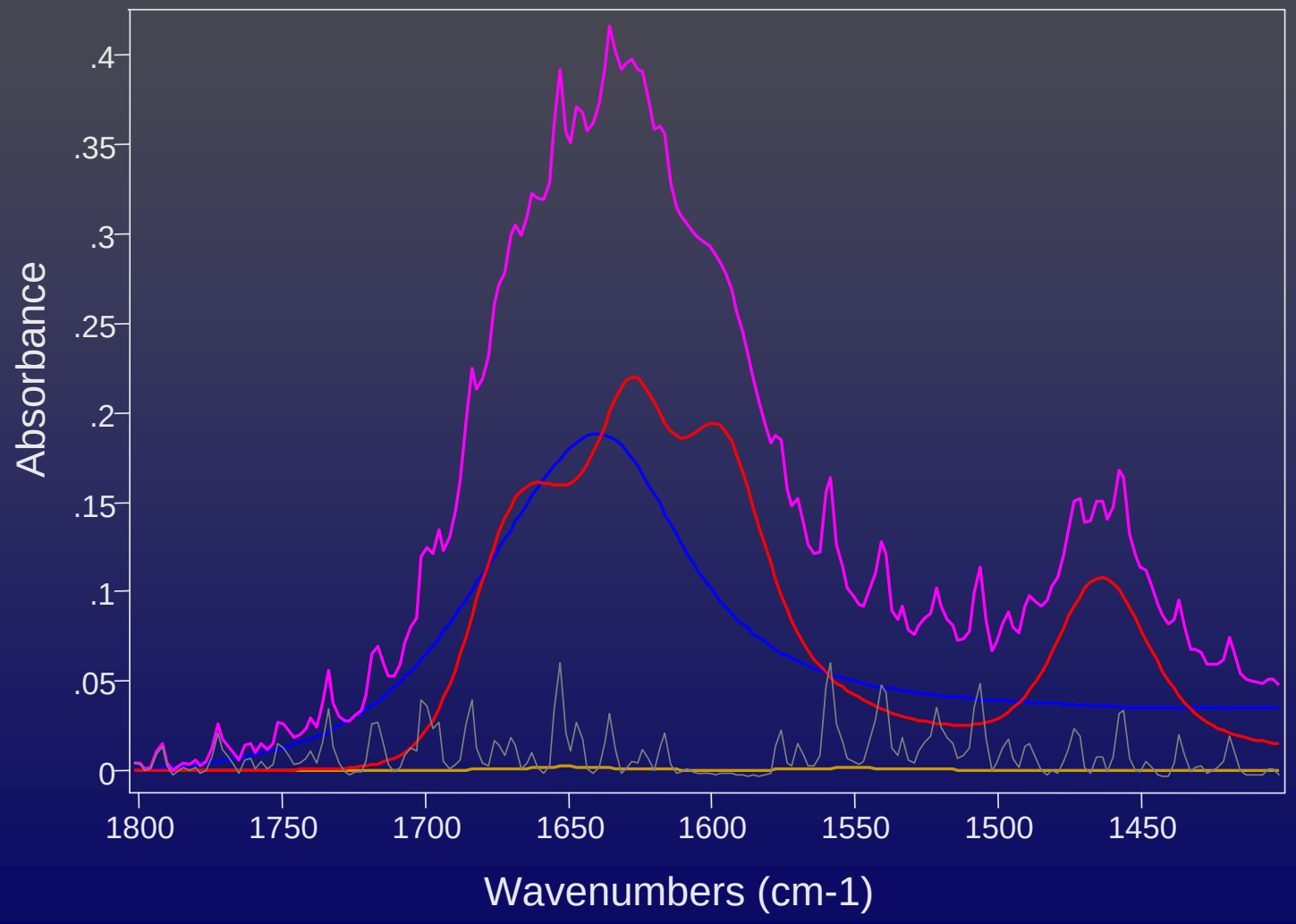
FIGURE 3: Interleukin-1β amide I curve-fit spectra from ATR-FTIR spectra showing the individual amide I components. IB is inclusion body, FA is folding aggregate, TA is thermal aggregate, and N is native. Reconstructed spectra are superimposed on original data; the fits were such that it is difficult to distinguish the two.

Difference Spectroscopy for Determining Structural Equivalence:

- Conformational comparison
- Stability comparison
- Identification of gross and subtle conformational differences
- Comparison of formulations

MEDICAL research PRODUCTS

FTIR Difference Spectroscopy



MEDICAL research PRODUCTS

FTIR Difference Spectroscopy

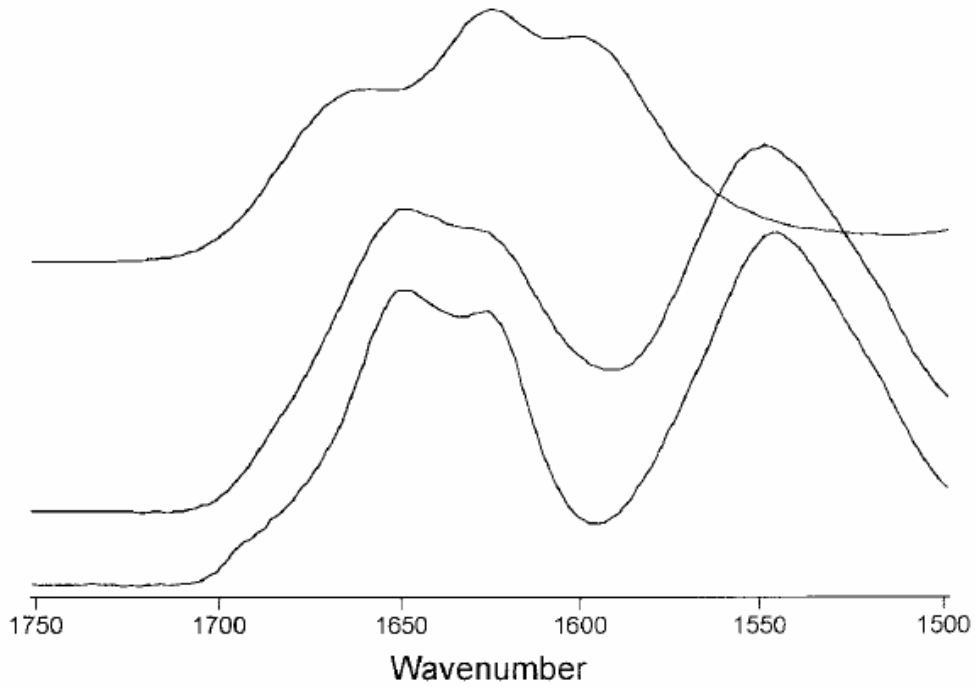


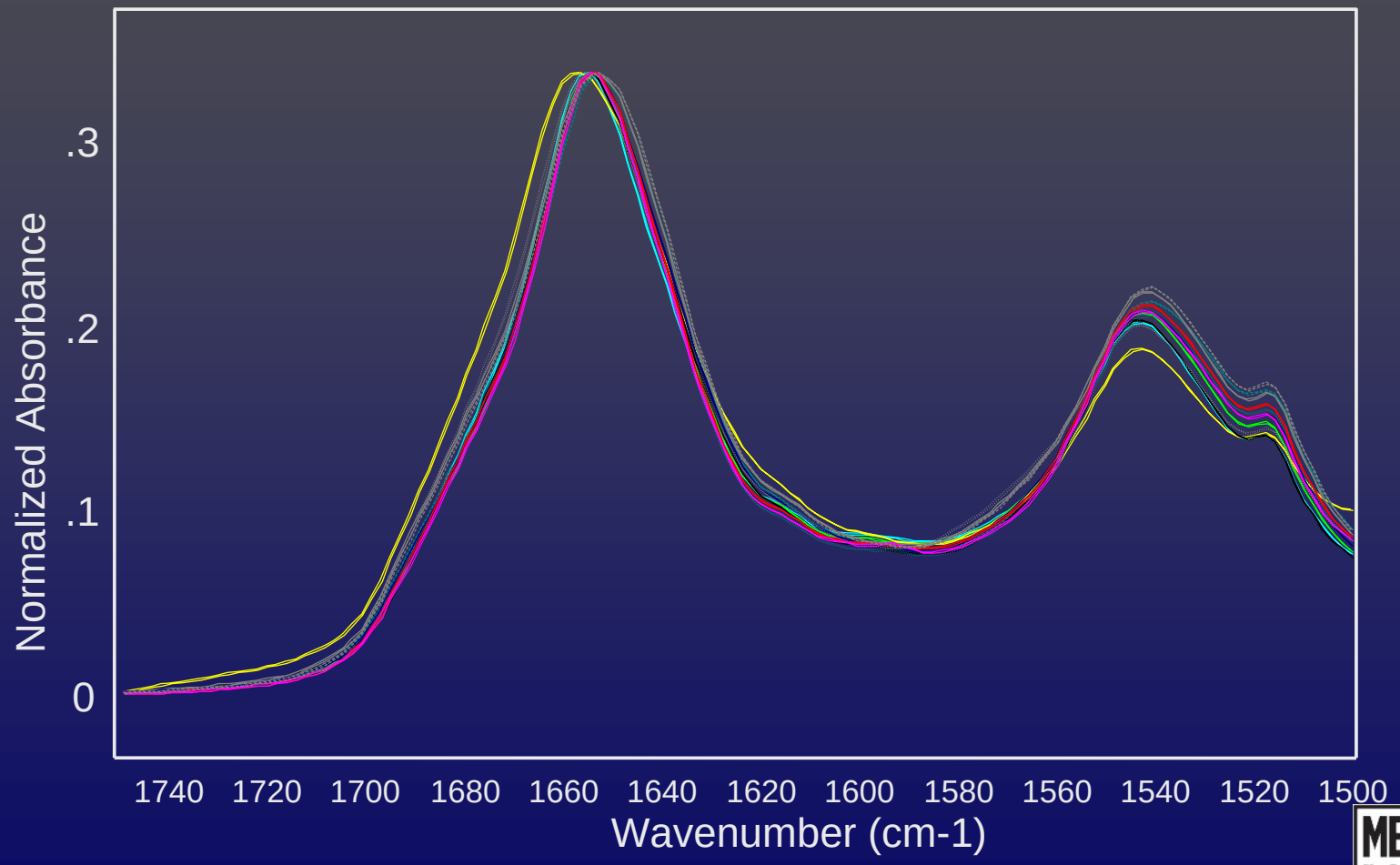
FIG. 11. Spectra of apomyoglobin with a "totally absorbing" cosolute, urea. The solutions used for this experiment would have been opaque in a transmission IR experiment. (Top) 8 M Urea with the water signal subtracted, (middle) protein signal from apomyoglobin in 8 M urea, (bottom) protein signal from extensively aggregated apomyoglobin in 2 M urea solution. The urea and H_2O signals have been subtracted from the apomyoglobin spectra.

Difference Enhancement Processing

- Useful in measuring formulation effects
- Second derivative overlap integral
  - Differentiate
  - Overlap integral
- Difference sharpening structural distortion score
  - Normalize
  - Generate difference spectrum
  - Fourier Self-Deconvolution
  - Integrate |result| = score

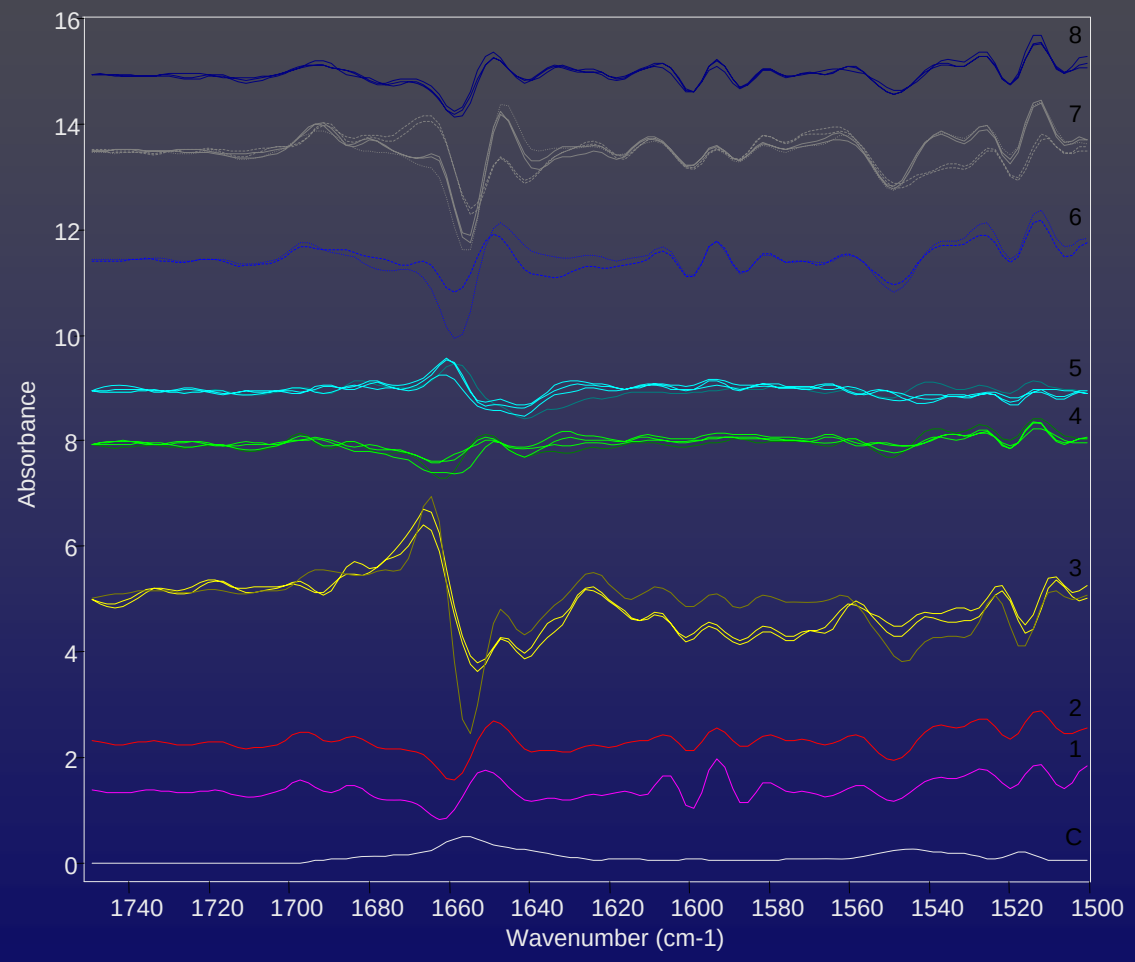
MEDICAL research PRODUCTS

Formulation Effects on Structure



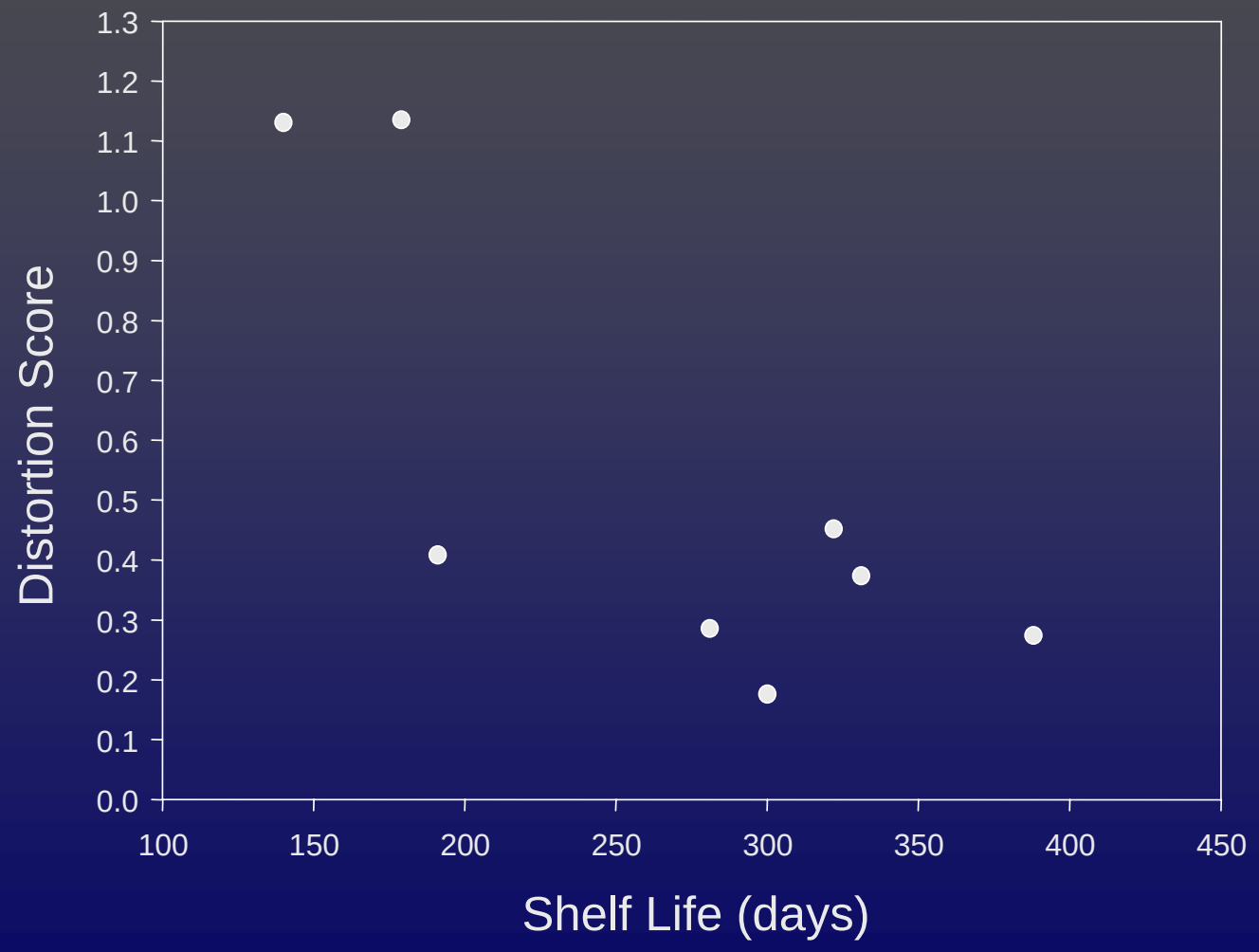
MEDICAL research PRODUCTS

Formulation Effects on Structure



Spectrum	Distortion Score
IF10816	5.19
IF20816	5.51
IF30816	40.37
IF30816B	34.93
IF3A0903	50.33
IF40816	4.14
IF4A0903	3.69
IF4B0903	1.51
IF4C0903	1.43
IF50816	2.82
IF5A0903	2.80
IF5B0903	2.06
IF5C0903	2.12
Controls	
IF60816P	15.58
IF60816S	6.50
IF7A0909	14.20
IF7B0909	14.35
IF7PE090	18.44
IF7S0909	13.06
IF7S2090	13.00
IF8A0909	5.95
IF8B0909	4.42
IF8C0909	4.75

Structural Distortion and Shelf Life



MEDICAL research PRODUCTS

Conclusions

- A broad range of low resolution optical spectroscopies are available that can probe a broad range of structure ensembles
- Optical techniques are rapid and reproducible, but do not stand alone
- Optical techniques can play role in the earliest stages of development
- Optical techniques are useful screens in troubleshooting development issues
- Optical spectroscopy can be useful on complex mixtures and physical states

MEDICAL research PRODUCTS

Acknowledgements

- Anthony Fink
- Ron Wetzel
- Margaret Speed
- Vladimir Uversky
- Erik Goormaghtigh
- Yvonne Looze

- Ron Lebel
- Peter Lord
- Hakan Edstrom
- Thomas Lobl
- Mark King

- John Rolf
- Xichen Zhang
- John Patterson

- Mark Hokenson
- Barbara Gorecka
- Sandy Strawbridge
- Sumukh Ray
- Punam Patel
- Vidya Gadamasetti
- Catalina Wong
- Alona Vizel

MEDICAL research PRODUCTS