Structure determination

NMR Structural Constraints

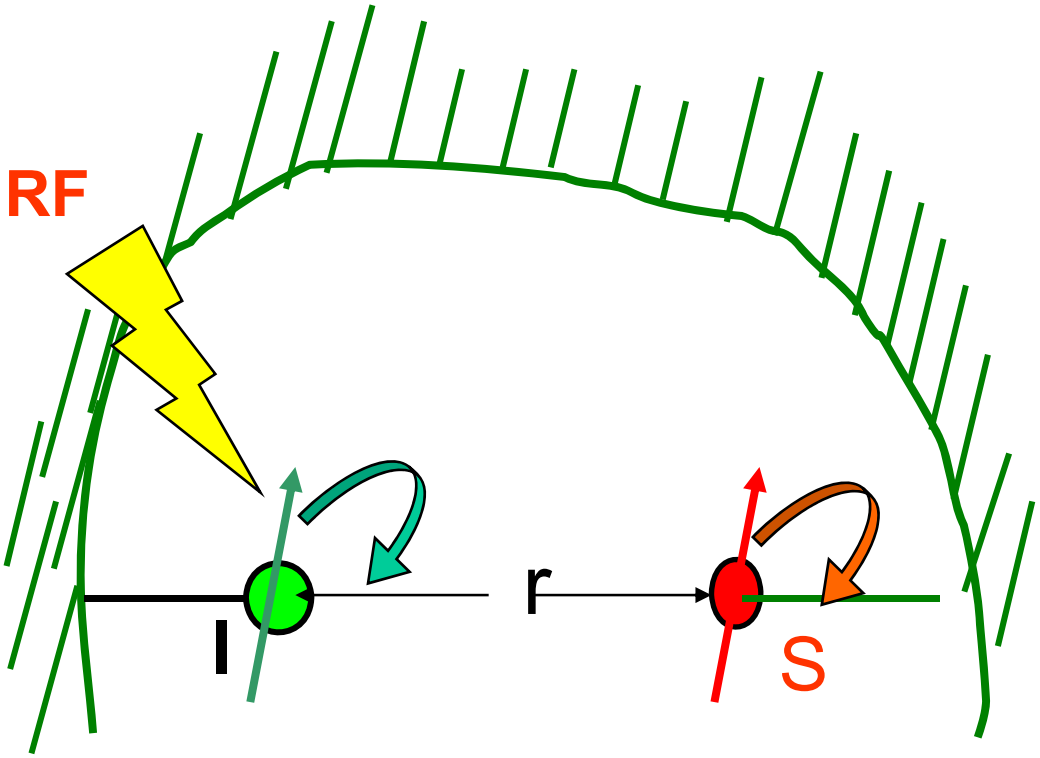
1. **Internuclear distances (Nuclear Overhauser Effect)**
   $NOE \propto R^{-6}$

2. **Dihedral angles (J-coupling):**
   $^{3}J_{NH\alpha} = 6.4 \cos^{2}(\Phi\text{-}60) - 1.4\cos(\Phi\text{-}60) + 1.9$

3. **Chemical Shift Index (CSI):**
   Chemical shift difference between observed and random coil chemical shift values → 2^{nd} structure determination

4. **Residual dipolar coupling:**
   Partial orientation of protein molecules in liquid crystal media permits observation of residual dipolar coupling for assessing long range orientations of dipolar coupled bonds.

3. Nuclear Overhauser Effect (NOE)



$$XNOE = 1 + (d^2/4)(\gamma_H/\gamma_N)[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]\, T_1$$

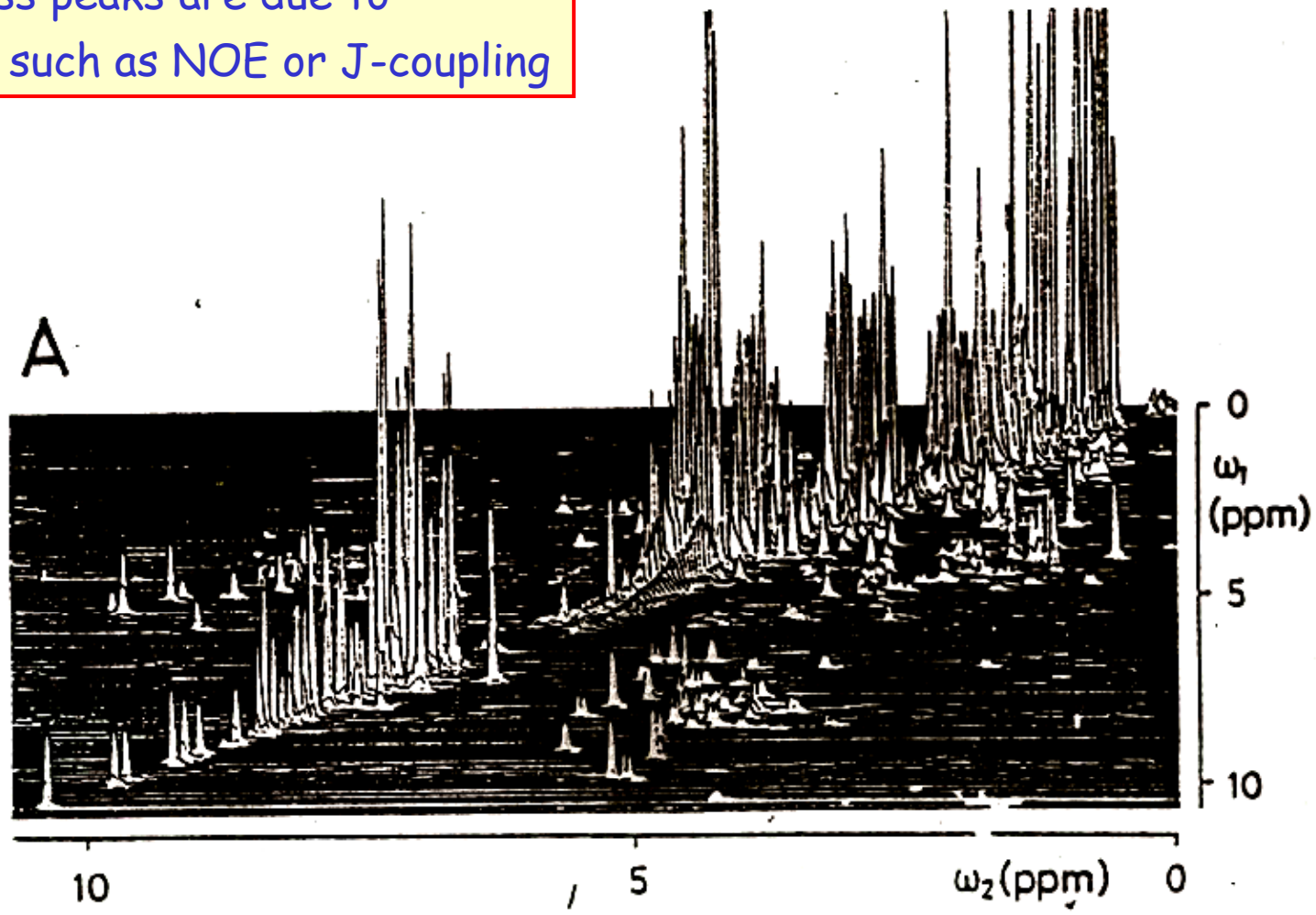
where $d = (\mu_o h \gamma_N \gamma_H / 8\pi^2)(r_{NH}^{-3})$, $J(\omega)$ is the spectral density function

➔ **1. Distance info: XNOE $\propto$ r^{-6} ;**

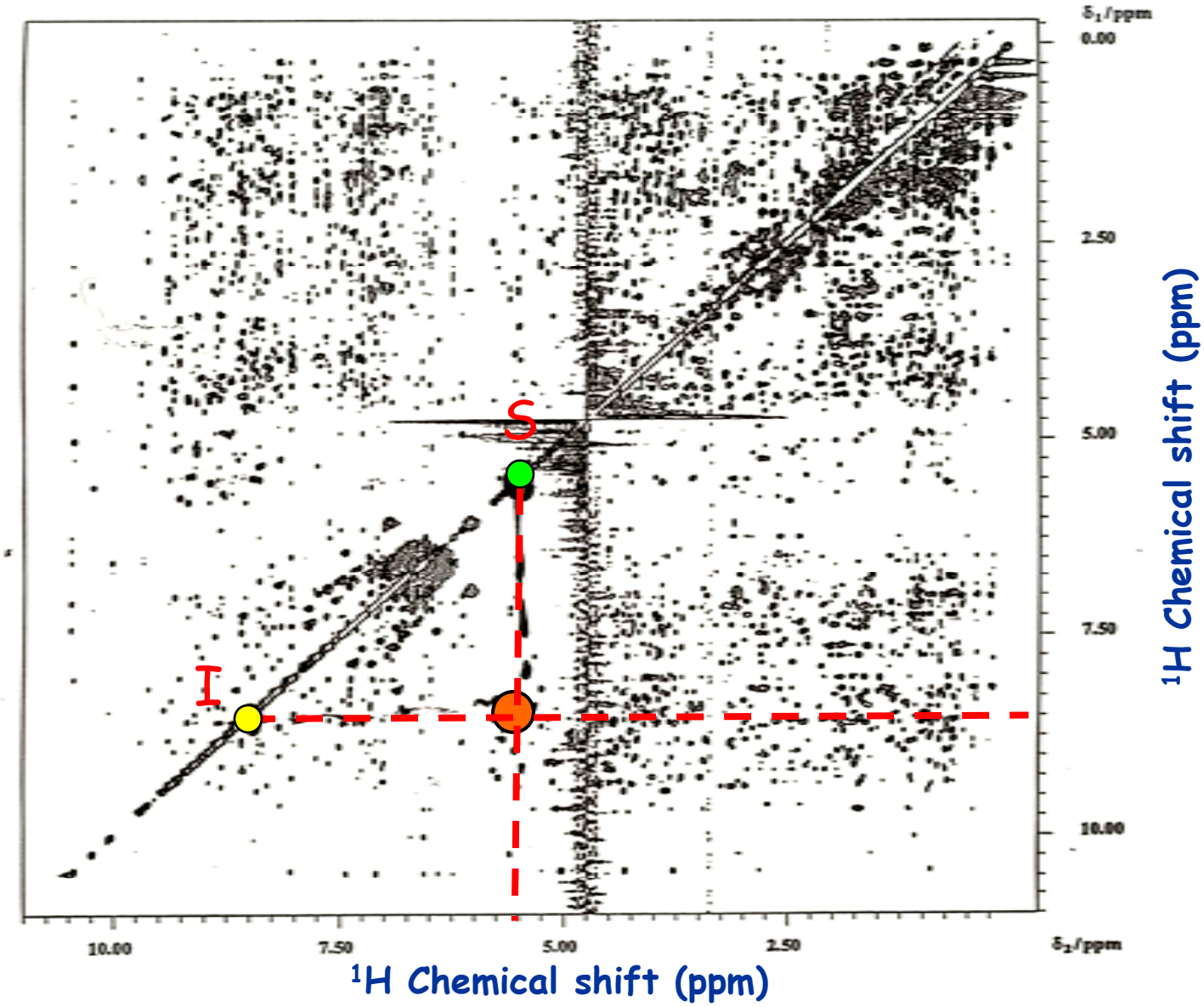
2. Dynamics: XNOE $\propto$ J (ω)

2D-NMR Spectrum

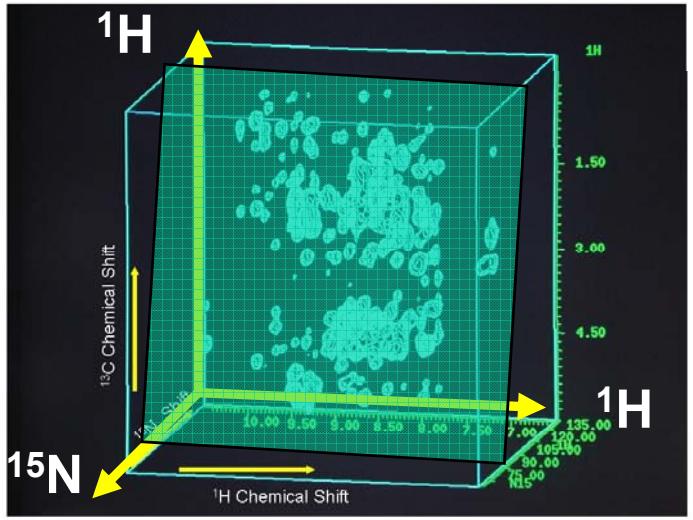
- Diagonal resonances same as in 1D spectrum
- off-diagonal cross peaks are due to **interactions** such as NOE or J-coupling

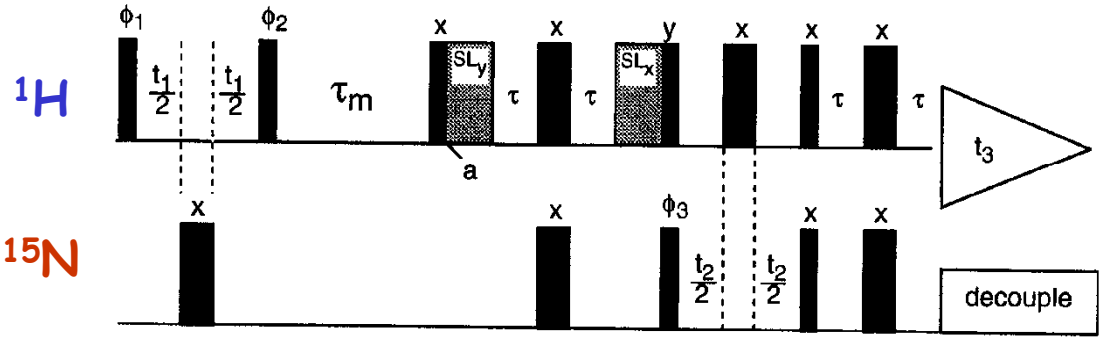


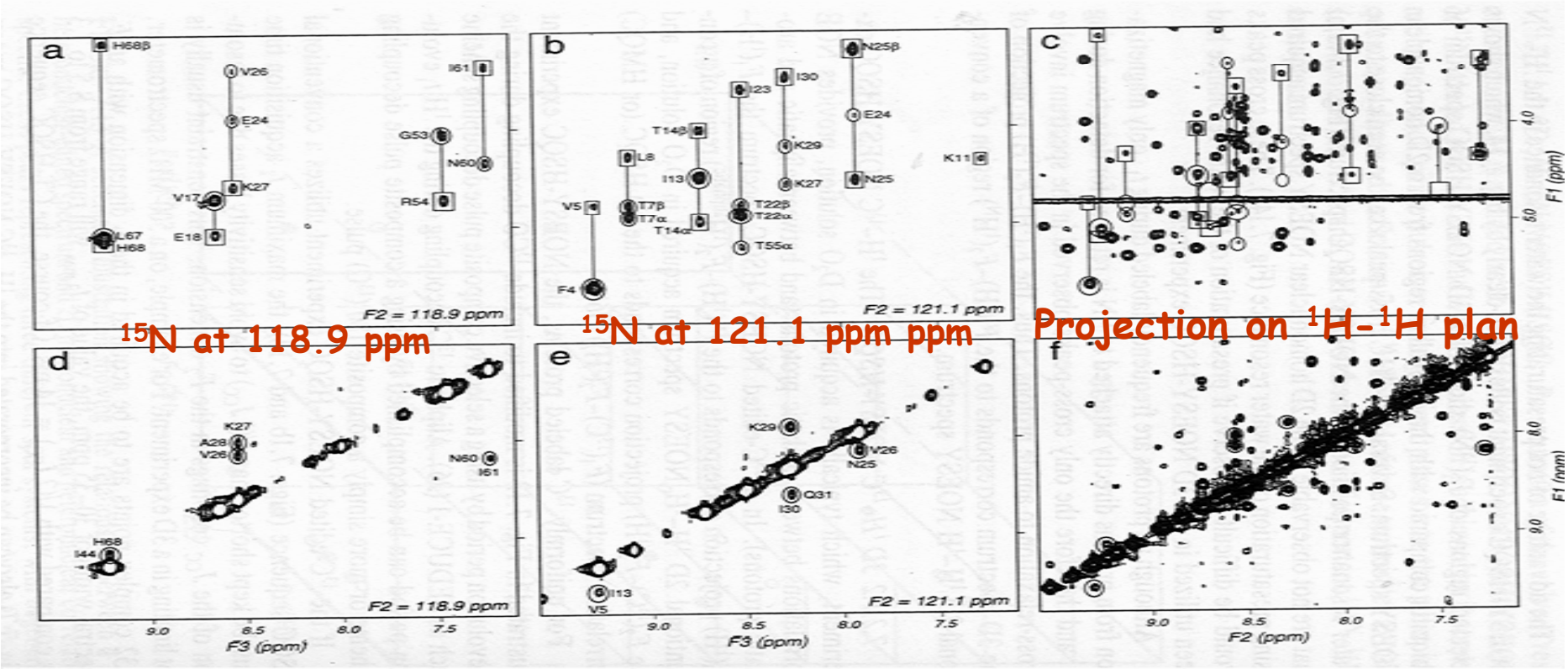
^{1}H - ^{1}H NOESY of RC-RNase



Heteronuclear-edited 3D NMR: ^{15}N-NOESY-HSQC





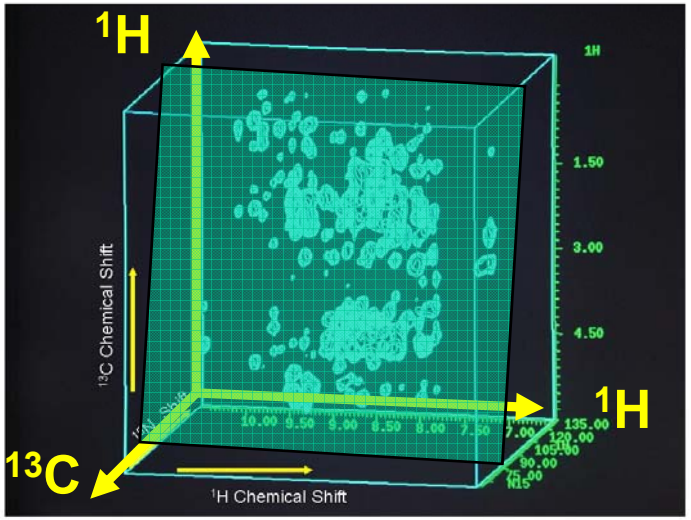


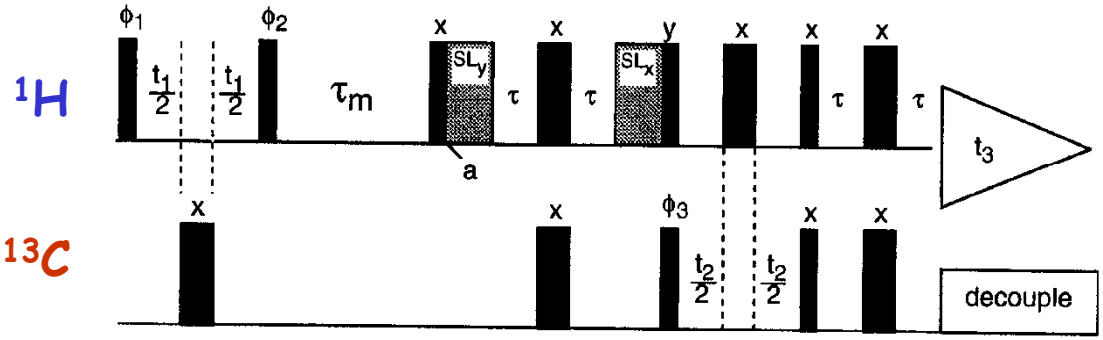
^{15}N at 118.9 ppm

^{15}N at 121.1 ppm ppm

Projection on ^{1}H-^{1}H plan

Heteronuclear-edited 3D NMR: 13C-NOESY-HSQC





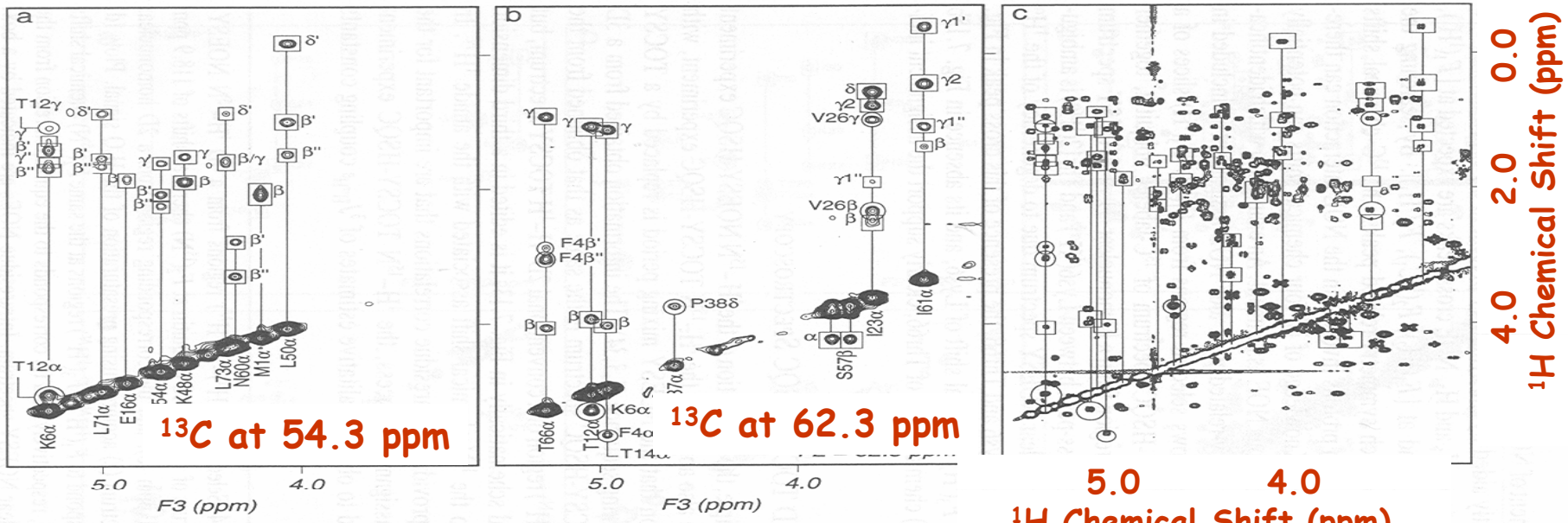
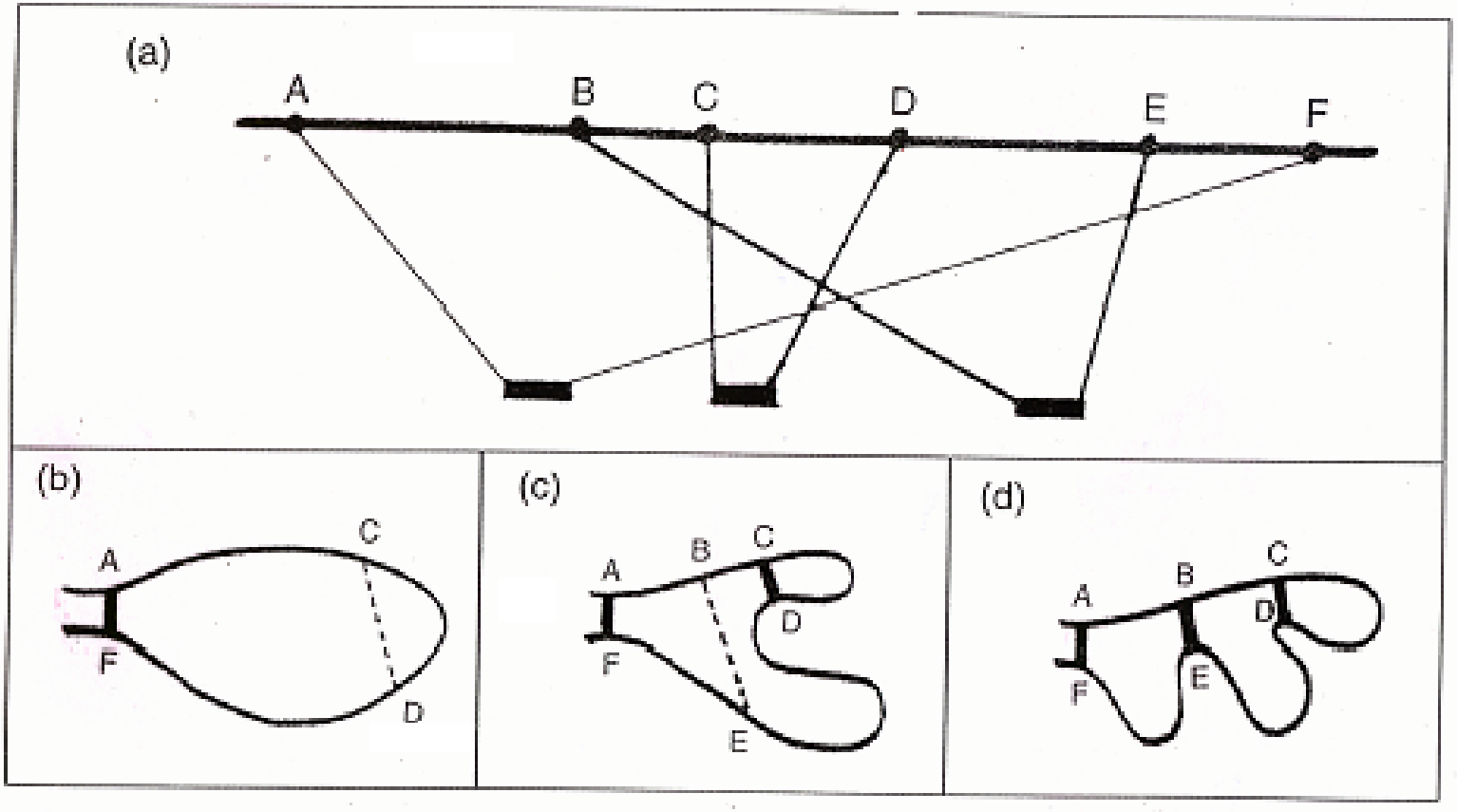


FIGURE 7.15 Selected $F_1(^1H)$–$F_3(^1H^\alpha)$ regions from a 3D ^1H–^{13}C NOESY- in in D_2O solution at $F_2(^{13}C)$ chemical shifts of 54.3 ppm (a) and 30.2/62.3 ppm (b), and the corresponding region from a 2D homonuclear NOESY spectrum (c) acquired using an unlabeled sample of ubiquitin in D_2O solution. Intraresidue NOEs are indicated by a box; interresidue NOEs are indicated by ellipses. The Lys6($^1H^\alpha$)–Thr12($^1H^\alpha$) cross-peak discussed in the text is located in the lower-left region of each spectrum.



由 NOE 限制找出分子的三度空間結構。

Medium range NOE of backbone resonances in secondary structures

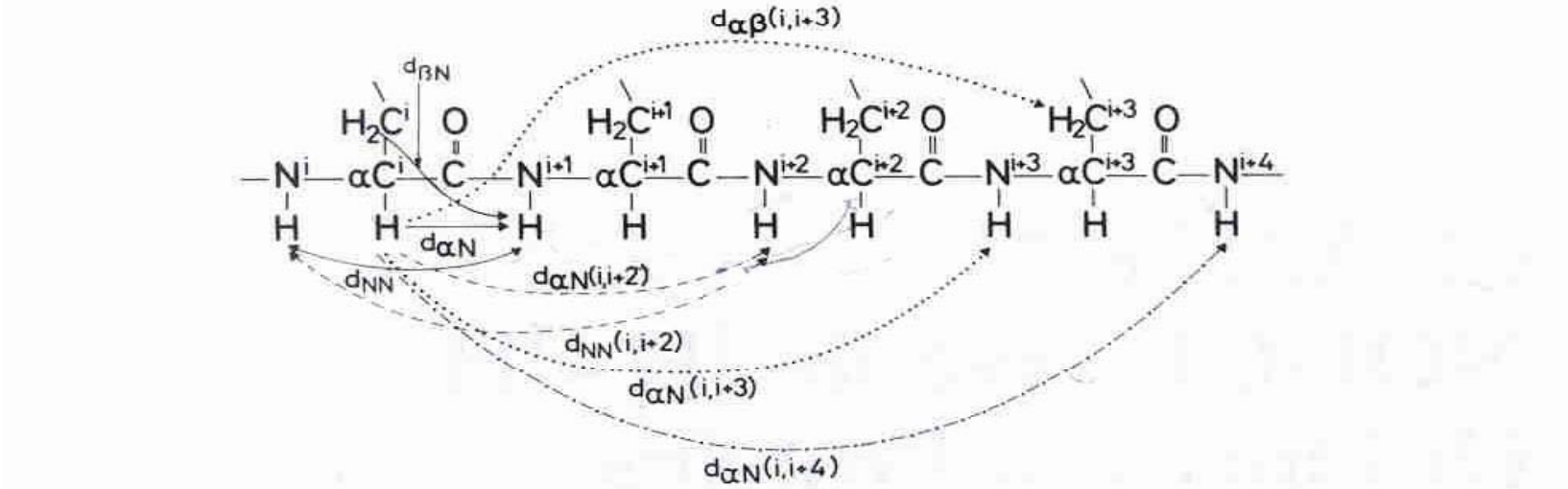
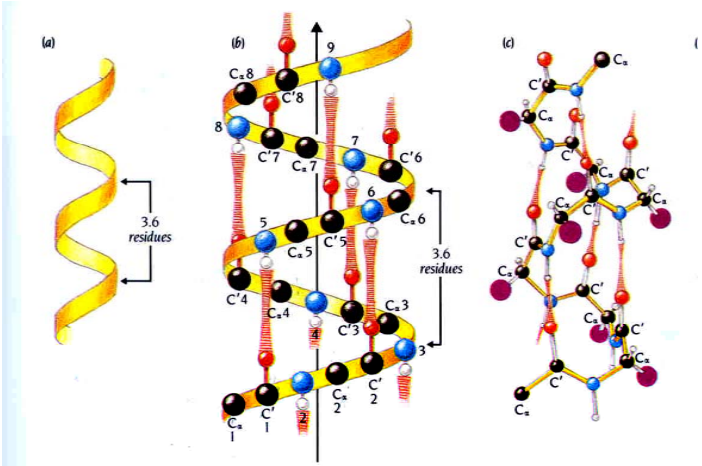
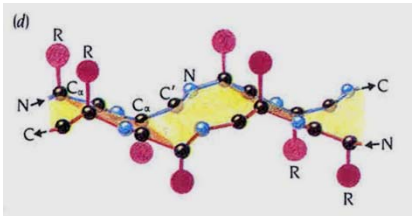


Figure 7.1. Selected sequential and medium-range $^1H-^1H$ distances in polypeptide chains (from Wüthrich et al., 1984a).

α-helix



β-Sheet



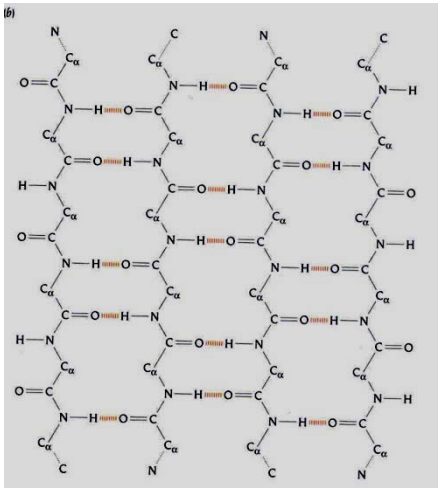


TABLE 7.1. Short (≤ 4.5 Å) Sequential and Medium-Range ^1H–^1H Distances in Polypeptide Secondary Structures

Distance	α-helix	3_{10}-helix	β	β_P	turn I[a]	turn II[a]
$d_{\alpha N}$	3.5	3.4	2.2	2.2	3.4	2.2
					3.2	3.2
$d_{\alpha N}(i,i+2)$	4.4	3.8			3.6	3.3
$d_{\alpha N}(i,i+3)$	3.4	3.3			3.1–4.2	3.8–4.7
$d_{\alpha N}(i,i+4)$	4.2					
d_{NN}	2.8	2.6	4.3	4.2	2.6	4.5
					2.4	2.4
$d_{NN}(i,i+2)$	4.2	4.1			3.8	4.3
$d_{\beta N}$[b]	2.5–4.1	2.9–4.4	3.2–4.5	3.7–4.7	2.9–4.4	3.6–4.6
					3.6–4.6	3.6–4.6
$d_{\alpha\beta}(i,i+3)$[b]	2.5–4.4	3.1–5.1				

[a] For the turns, the first of two numbers applies to the distance between residues 2 and 3, the second to that between residues 3 and 4 (Fig. 7.12). The range indicated for $d_{\alpha N}(i,i+3)$ corresponds to the distances adopted if ψ_1 is varied between −180 and 180°.

[b] The ranges given correspond to the distances adopted by a β-methine proton if χ^1 is varied between −180 and 180°.

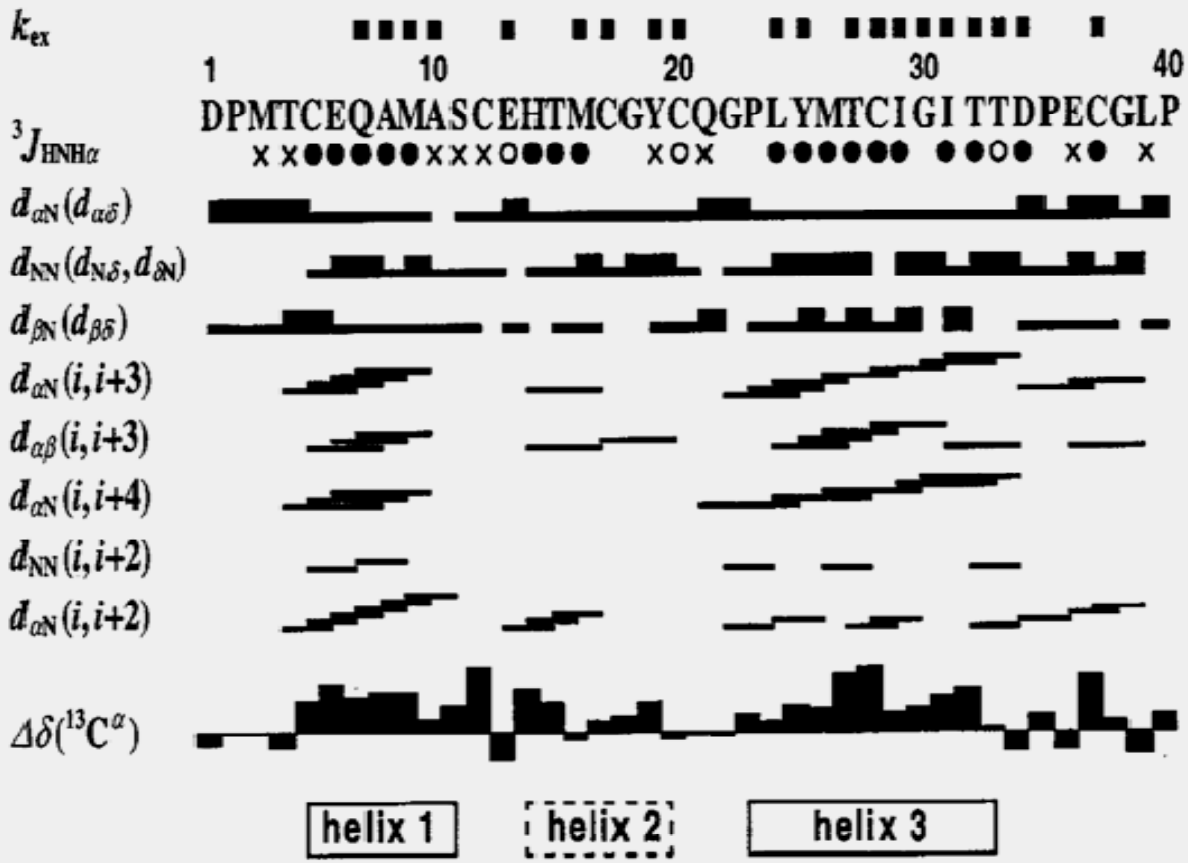


Figure 8. Example showing methods recommended for presenting NMR data supporting the identification of regular secondary structure in proteins. The 40-residue protein, pheromone Er-2, is used as an illustration (ref. 51). Above the amino acid sequence, black squares identify residues with observably slow hydrogen-exchange rates, k_{ex}, at the backbone amide (the conditions of the exchange experiment should be specified). Below the amino acid sequence, filled circles identify residues with $^3J_{HNH\alpha} < 6.0$ Hz, indicative of local α-type conformation; open circles correspond to $^3J_{HNH\alpha} > 8.0$ Hz, indicative of residues in extended chain conformation; crosses identify residues with $^3J_{HNH\alpha}$ values 6.0 to 8.0 Hz. For the sequential proton–proton NOE connectivities, $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ ($d_{\alpha\delta}$, $d_{N\delta}$, and $d_{\beta\delta}$ for Xxx–Pro dipeptides, $d_{\alpha N}$, $d_{\delta N}$, $d_{\beta N}$ for Pro–Xxx dipeptides), thick and thin bars indicate strong and weak NOE intensities, respectively. The observed medium-range NOEs $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$, $d_{\alpha N}(i,i+4)$, $d_{NN}(i,i+2)$, and $d_{\alpha N}(i,i+2)$ are indicated by lines connecting the two residues that are related by the NOE. $^{13}C^{\alpha}$ chemical shifts relative to the random coil values, $\Delta\delta(^{13}C^{\alpha})$, are plotted at the bottom of the Figure, where positive values are shifts to lower field. The sequence locations of three helices are indicated at the bottom; broken lines are used to indicate that the identification of helix 2 from these data is uncertain.

Use of J-coupling for structure determination

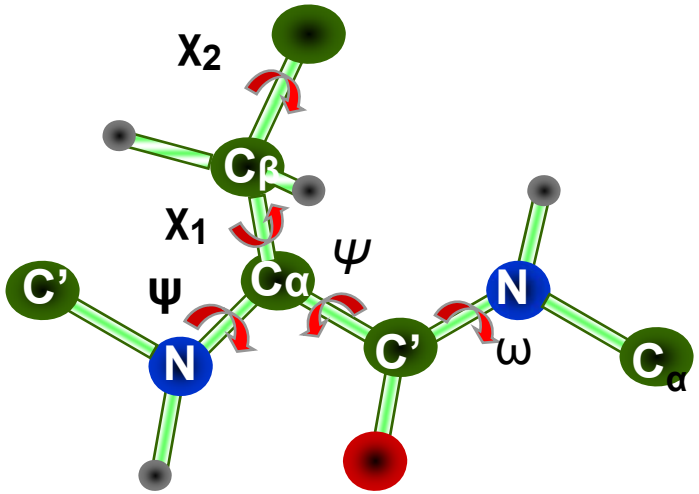
- **One neighboring spins: - $C_\alpha H$ - NH -**

- **From *coupling constant* (J) one can determine the dihedral angles from the following Karplus equations, where $^3J_{NH\alpha}$ is the coupling constant between $C_\alpha H$ - NH.**

$$^3J_{NH\,\alpha} = 6.4\cos^2(\phi - 60) - 1.4\cos(\phi - 60) + 1.9$$

$$^3J_{\alpha\beta\,1} = 9.5\cos^2(\chi_1 - 120) - 1.6\cos(\chi_1 - 120) + 1.8$$

$$^3J_{\alpha\beta\,2} = 9.5\cos^2\chi_1 - 1.6\cos\chi_1 + 1.8$$



- $^3J_{NH\alpha}$ = 4 – 11 Hz depends on secondary structure.
  $^3J_{NH\alpha}$ < 6 Hz → α-helix; $^3J_{NH\alpha}$ > 8 Hz → β-stand

Chemical Shift Indices (CSI): ($^1H_\alpha$, $^{13}C_\alpha$, $^{13}C_\beta$, ^{13}CO)

CSI is a commonly accepted procedure to establish the secondary structure of proteins based on chemical shift differences with respect to some predefined 'random coil' values (Secondary shift). It can be applied from the measured HA, CA, CB and CO chemical shifts for each residue in a protein.

Secondary shift: The difference between the observed shift and the corresponding random coil value.

- Valuable in identifying secondary structure, determining ring pucker, delineating flexible regions, locating hydrogen bonds, setting dihedral restraints, and detecting aromatic stacking interactions.

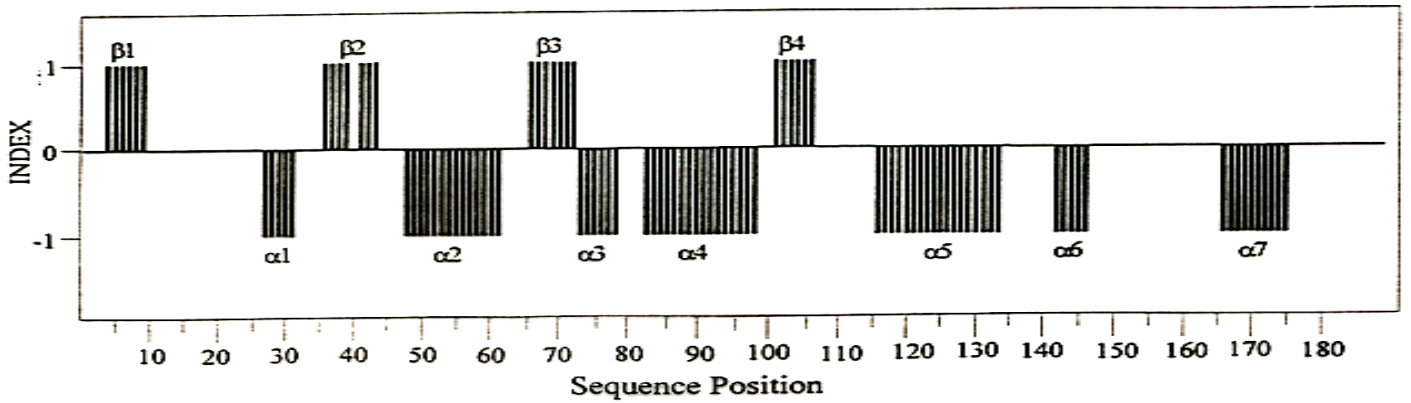


Figure 5. CSI consensus plot for *E. coli* thioesterase/protease I, determined using four nuclei ($^1H^\alpha$, $^{13}C^\alpha$, $^{13}C^\beta$, and ^{13}CO). The secondary structural motifs obtained from this program are summarized in the figure.

Chemical Shift Referencing: The ^{1}H chemical shift was referenced to 2,2-dimethyl-2-Silapentane-5-sulfonate (DSS) at 0 ppm. The ^{15}N and ^{13}C chemical shift values were referenced using the consensus ratio of Ξ of 0.101329118 and 0.251449530 for $^{15}N/^{1}H$ and $^{13}C/^{1}H$, respectively

(Wishart and Case, Method. Enzymol. 338, 3-34 (2001))

TABLE I

IUPAC/IUBMB RECOMMENDED Ξ (XI) RATIOS FOR INDIRECT CHEMICAL SHIFT REFERENCING IN BIOMOLECULAR NMR[a]

Nucleus	Compound	Ξ Ratio
^{1}H	DSS	1.000 000 000
^{13}C	DSS	0.251 449 530
^{15}N	Liquid NH_3	0.101 329 118
^{19}F	CF_3COOH	0.940 867 196
^{31}P	$(CH_3)_3PO_4$	0.404 808 636

[a] Relative to DSS.

Ξ ratio (Nucleus-specific frequency ratio: Determine the precise ^{1}H resonance frequency of DSS then multiply this frequency by Ξ of a particular nucleus one obtains the exact resonance frequency reference at 0 ppm of that nucleus.

Random Coil Chemical Shift Table

Residue	^{1}HN	^{15}N	HA	CA	CB	CO
Ala	8.28	113.5	4.35	52.5	19.0	177.1
Cys	8.32	118.8	4.65	58.8	28.6	174.8
Asp	8.34	120.4	4.76	54.1	40.8	177.2
Glu	8.42	120.2	4.29	56.7	29.7	176.1
Phe	8.30	120.3	4.66	57.9	39.3	175.8
Gly	8.33	108.8	3.97	45.0	-	173.6
His	8.42	118.2	4.63	55.8	32.0	175.1
Ile	8.00	119.9	3.95	62.6	37.5	176.8
Lys	8.29	120.4	4.36	56.7	32.3	176.5
Leu	8.16	121.8	4.17	55.7	41.9	177.1
Met	8.28	119.6	4.52	56.6	32.8	175.5
Asn	8.40	118.7	4.75	53.6	39.0	175.5
Pro			4.44	62.9	31.7	176.3
Gln	8.32	119.8	4.37	56.2	30.1	176.0
Arg	8.23	120.5	4.38	56.3	30.3	176.5
Ser	8.31	115.7	4.50	58.3	62.7	173.7
Thr	8.15	113.6	4.35	63.1	68.1`	175.2
Val	8.03	119.2	3.95	63.0	31.7	177.1
Trp	8.25	121.3	4.70	57.8	28.3	175.8
Tyr	8.25	113.3	4.60	58.6	38.7	175.7

Wishart and Nip, Biochem. Cell Biol. Vol. 76, 1998

TABLE 2.3. Random Coil ^1H Chemical Shifts for the 20 Common Amino Acid Residues[a]

Residue	NH	αH	βH	Others
Gly	8.39	3.97		
Ala	8.25	4.35	1.39	
Val	8.44	4.18	2.13	γCH_3 0.97, 0.94
Ile	8.19	4.23	1.90	γCH_2 1.48, 1.19
				γCH_3 0.95
				δCH_3 0.89
Leu	8.42	4.38	1.65,1.65	γH 1.64
				δCH_3 0.94, 0.90
Pro[b]		4.44	2.28,2.02	γCH_2 2.03, 2.03
				δCH_2 3.68, 3.65
Ser	8.38	4.50	3.88,3.88	
Thr	8.24	4.35	4.22	γCH_3 1.23
Asp	8.41	4.76	2.84,2.75	
Glu	8.37	4.29	2.09,1.97	γCH_2 2.31, 2.28
Lys	8.41	4.36	1.85,1.76	γCH_2 1.45, 1.45
				δCH_2 1.70, 1.70
				εCH_2 3.02, 3.02
				εNH_3^+ 7.52
Arg	8.27	4.38	1.89,1.79	γCH_2 1.70, 1.70
				δCH_2 3.32, 3.32
				NH 7.17, 6.62
Asn	8.75	4.75	2.83,2.75	γNH_2 7.59, 6.91
Gln	8.41	4.37	2.13,2.01	γCH_2 2.38, 2.38
				δNH_2 6.87, 7.59
Met	8.42	4.52	2.15,2.01	γCH_2 2.64, 2.64
				εCH_3 2.13
Cys	8.31	4.69	3.28,2.96	
Trp	8.09	4.70	3.32,3.19	2H 7.24
				4H 7.65
				5H 7.17
				6H 7.24
				7H 7.50
				NH 10.22
Phe	8.23	4.66	3.22,2.99	2,6H 7.30
				3,5H 7.39
				4H 7.34
Tyr	8.18	4.60	3.13,2.92	2,6H 7.15
				3,5H 6.86
His	8.41	4.63	3.26,3.20	2H 8.12
				4H 7.14

[a] Data for the nonterminal residues X in tetrapeptides GGXA, pH 7.0, 35°C [from Bundi and Wüthrich (1979a), except that more precise data were obtained for Leu, Pro, Lys, Arg, Met, and Phe using new measurements at 500 MHz].
[b] Data for *trans*-Pro.

Chemical Shift Index (CSI) for Secondary Structure Determination

Wishart et al. Meth. Enzymol. 338, 3-34 (2002)
Wishart and Sykes, J Biomol NMR 4, 171 (1994)

Simple rules

"1" if the measured chemical shift is greater than the specific CSI value. "-1" if it's smaller. "0" if it within the expected range.

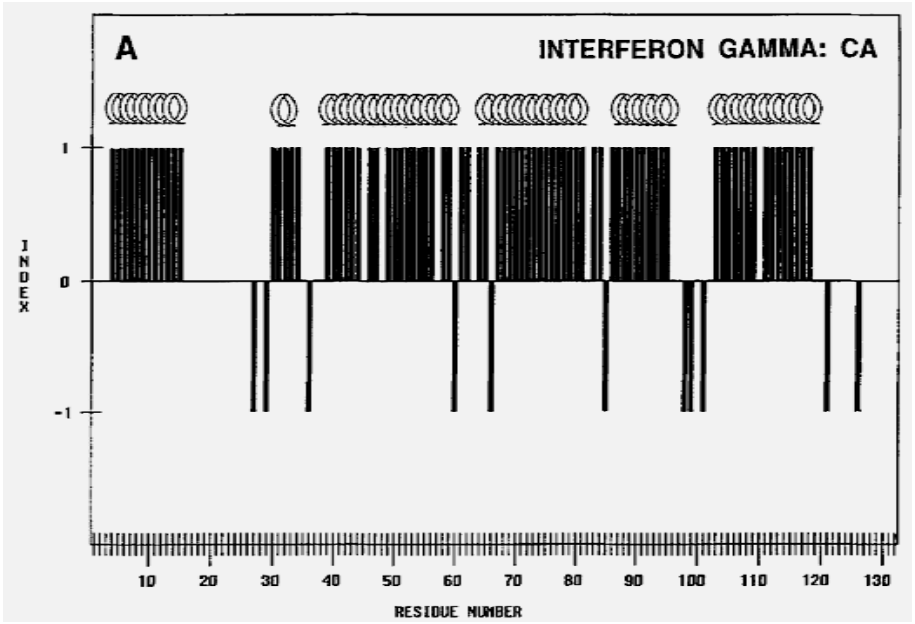
Limits on allowable deviations from random coil shifts :

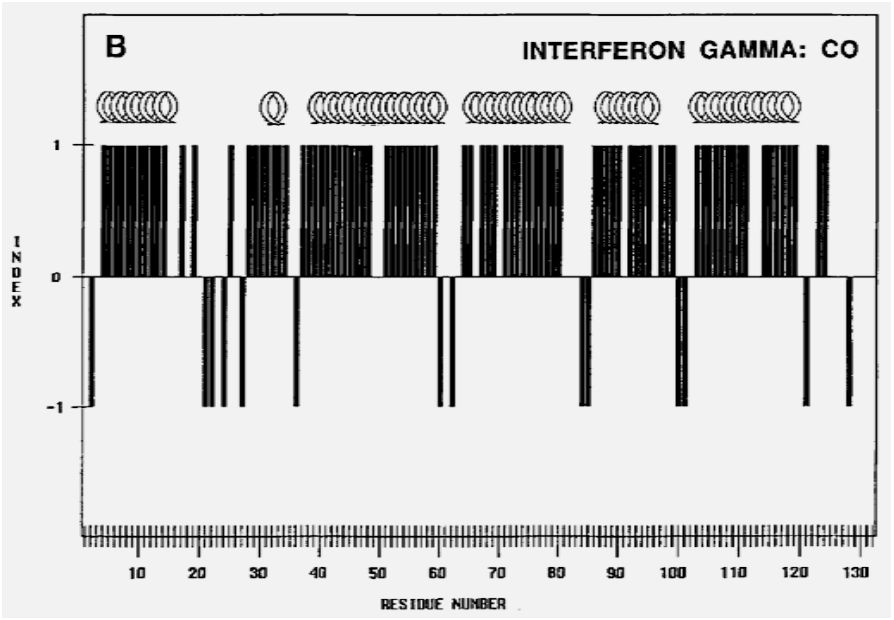
$^{1}H\alpha$: ±1.3 ppm; **$^{13}C\alpha$ and $^{13}C\beta$: ±5 ppm;**

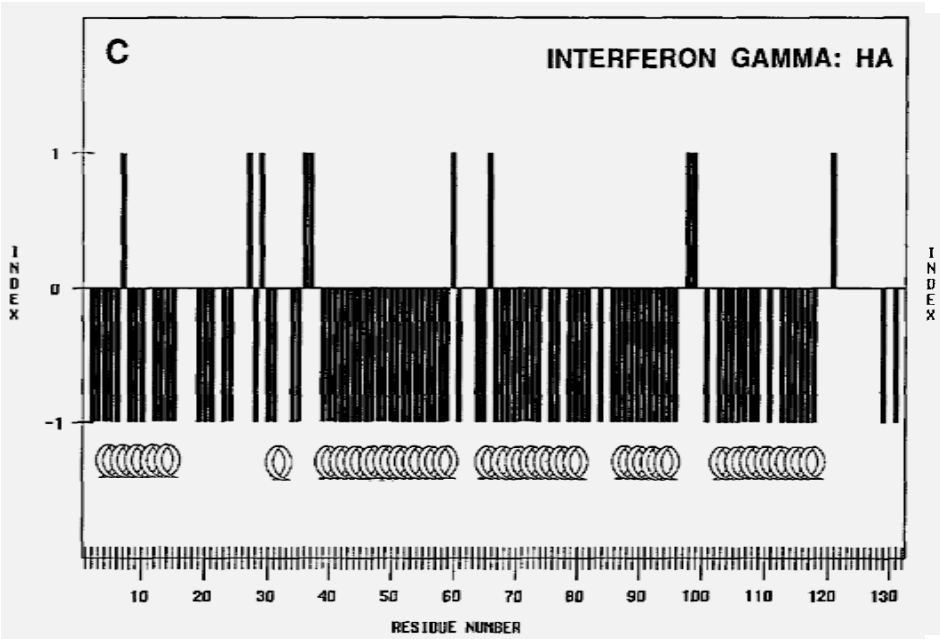
^{13}CO: ±4 ppm **^{15}N: ± 10 ppm**

Assigning the secondary structure

- Alpha helix is defined when four or more "-1" HA and/or "1" CA/CO are sequentially found.
- A beta-strand is defined when three or more "1" HA and/or "-1" CA/CO are sequentially found.
- All other regions are designated as coil.
- When several CSIs are available, a **consensus CSI** based on the majority rule approach allows to improve the prediction of the secondary structure.


A
INTERFERON GAMMA: CA
INDEX
1
0
-1
10 20 30 40 50 60 70 80 90 100 110 120 130
RESIDUE NUMBER


B
INTERFERON GAMMA: CO
INDEX
1
0
-1
10 20 30 40 50 60 70 80 90 100 110 120 130
RESIDUE NUMBER


C
INTERFERON GAMMA: HA
INDEX
1
0
-1
10 20 30 40 50 60 70 80 90 100 110 120 130
RESIDUE NUMBER

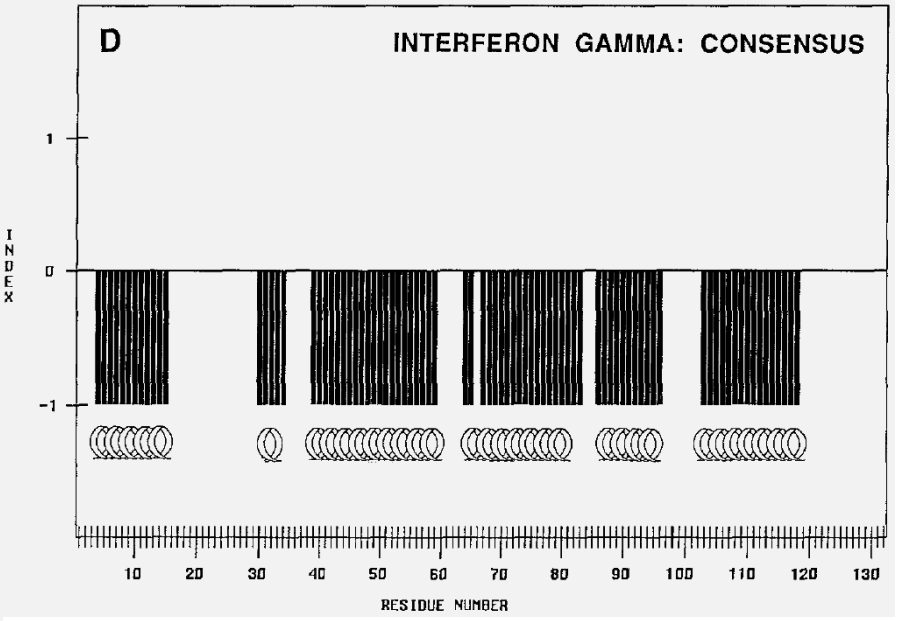

D
INTERFERON GAMMA: CONSENSUS
INDEX
1
0
-1
10 20 30 40 50 60 70 80 90 100 110 120 130
RESIDUE NUMBER

TABLE 3
ACCURACY OF CHEMICAL-SHIFT INDICES (%) FOR 20 FULLY ASSIGNED PROTEINS

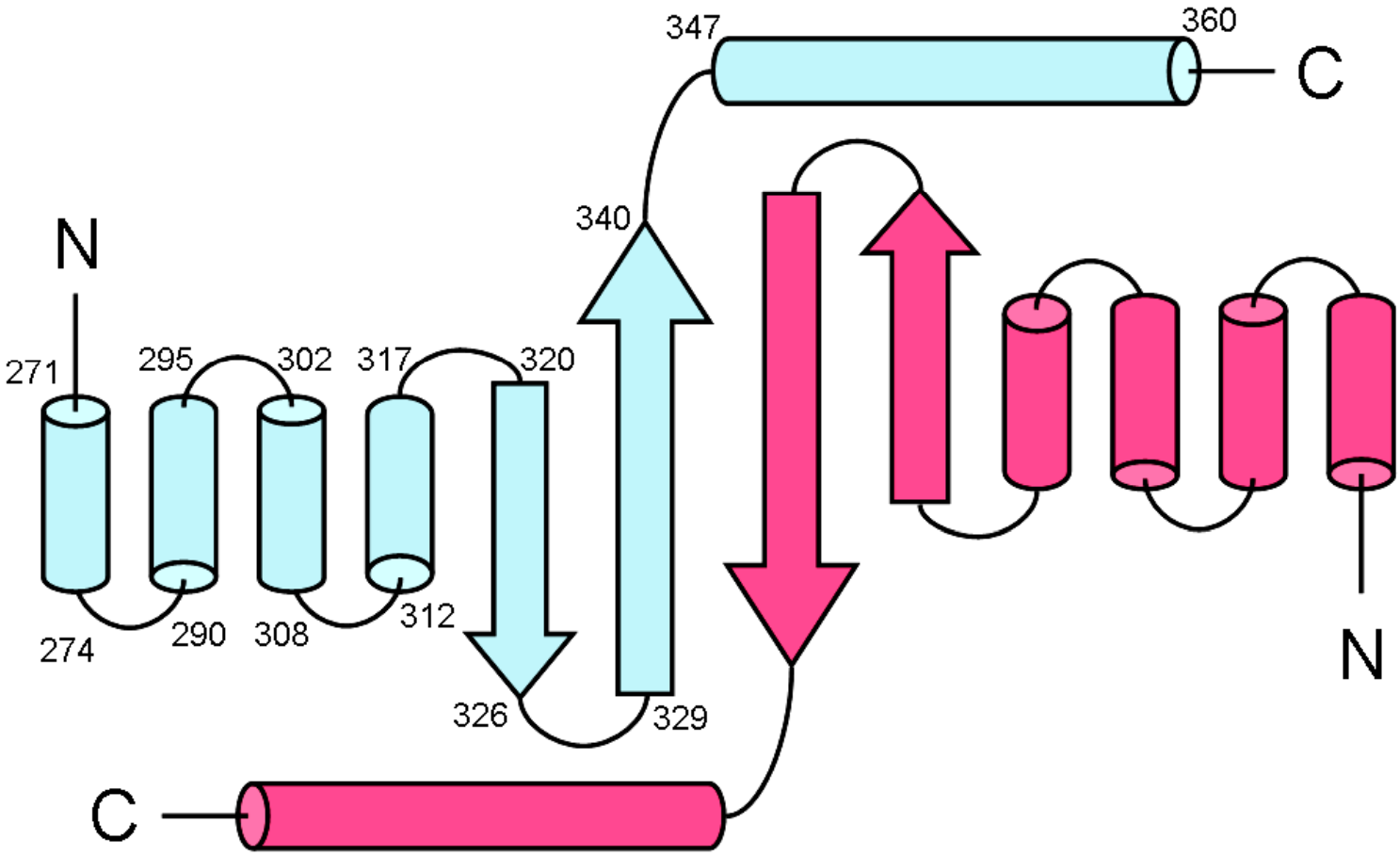
Protein (reference)	$^{13}C^{\alpha}$ CSI	$^{13}C'$ CSI	$^{13}C^{\beta}$ CSI	$^{1}H^{\alpha}$ CSI	Consensus (%)
Gal 4 (Shirakawa et al., 1993)	98	–[a]	85	90	100
Troponin C* (C. Slupsky)[b]	93	91	–	91	97
Interferon gamma* (Grzesiek et al., 1992)	92	88	–	89	94
Calmodulin* (Ikura et al., 1990, 1991)	91	76	98	82	94
Ribonuclease H* (Yamazaki et al., 1991, 1993)	86	75	94	85	94
Interleukin 4* (Powers et al., 1992)	90	93	99	90	93
hn RNP c* (Wittekind et al., 1992)	88	89	–	84	93
SH2 Domain-plγ (J. Forman-Kay)[b]	88	–	83	83	93
FKB 506 (Xu et al., 1993)	88	74	88	89	93
Digoxin antibody (Constantine et al., 1993)	88	–	87	87	92
Tendamistat* (Kessler et al., 1990)	85	–	78	85	91
Glucose Perm. IIA* (Fairbrother et al., 1992)	82	–	83	81	90
Cellulose BP (L. Kay)[b]	83	77	85	76	90
Profilin (Archer et al., 1993)	84	79	81	81	90
Staph. Nuclease (T. Yamazaki)[b]	81	70	84	87	87
Phosphocarrier III* (Pelton et al., 1991)	82	76	82	75	86
Ras P-21 (Campbell-Burk et al., 1992)	83	–	–	86	–
Interleukin 1β* (Clore et al., 1990)	85	–	–	84	–
IRAP* (Stockman et al., 1992)	82	–	–	81	–
BPTI* (Wagner and Bruhwiler, 1986)	78	–	–	78	–

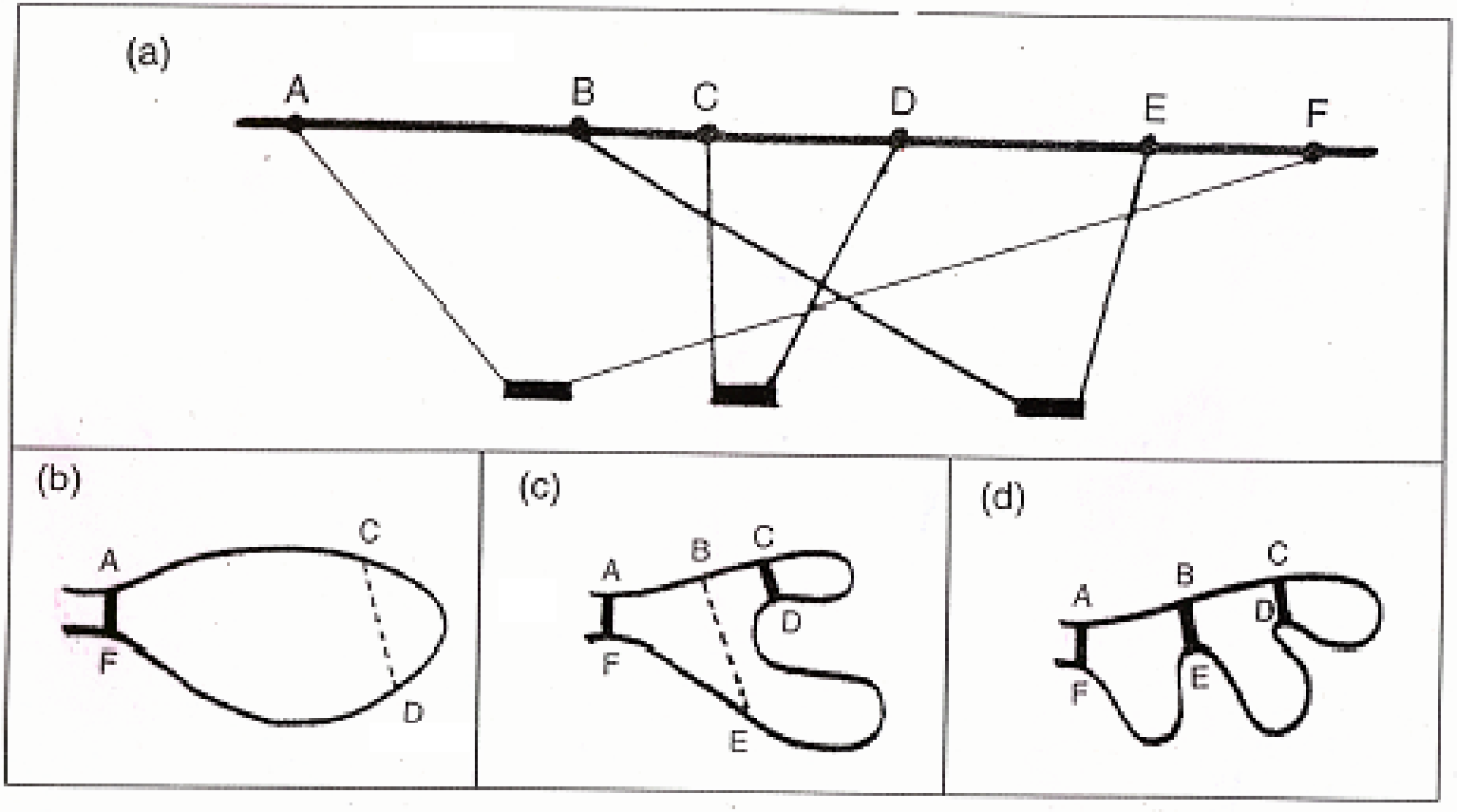
* Indicates that the chemical shifts from these proteins were used in compiling the preliminary reference set of CSI values.
[a] Chemical-shift data unavailable or not published.
[b] Personal communication.

Stephan Schwarzinger, Gerard J. A. Kroon, Ted R. Foss, John Chung, Peter E. Wright,* and H. Jane Dyson* *J. Am. Chem. Soc.* 2001, *123,* 2970-2978 "Sequence-Dependent Correction of Random Coil NMR Chemical Shifts"

Tertiary structure (Backbone rmsd: 1.72Å/1.0Å)

(CYANA with extensive manual assignments)

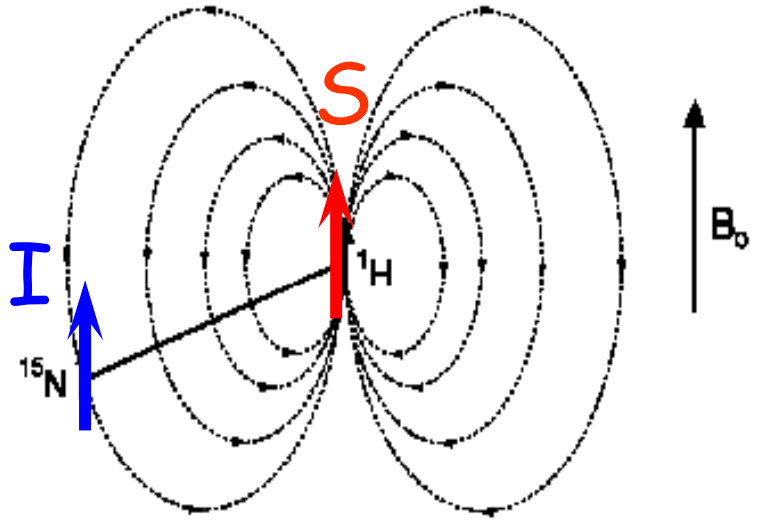




由 NOE 限制找出分子的三度空間結構。

Refinement with residual dipolar coupling (RDC)

Dipolar Field



Residual Dipolar Coupling:

S: Scaling factor = 0 in non-viscous liquid media

γ: Nuclear gyric ratios

$$D_{IS}(\theta,\varphi) = -\left(\frac{\mu_o h S \gamma_I \gamma_S}{16\pi R^3}\right)\left\{(3\cos^2\theta - 1) + \frac{3}{2}A\sin^2\theta\cos 2\varphi\right\}$$

R: Distance between spins I and S

Structure Calculation

1. Build a random structure of the given sequence.
2. Use molecular dynamics and simulated annealing to generate many structures with minimum violation of structure constraints and with minimal energy of the following energy term.

   $E_{total} = E_{bond} + E_{angle} + E_{improper} + E_{VDW} + E_{cdih} + E_{NOE} + E_{RDC} + \ldots.$

   $E_{bond} = \Sigma k_b(b-b_0)^2;$ $E_{\varphi} = \Sigma k_{\varphi}(\varphi - \varphi_0)^2;$

   $E_{improper} = \Sigma k_{impr}(\omega - \omega_0)^2;$ $E_{cdih} = \Sigma k_{cdih}(\Psi - \Psi_0)^2;$

   $E_{NOE} = \Sigma k_{NOE}(\gamma - \gamma_0)^2;$ $E_{RDC} = \Sigma k_{RDC}(\theta - \theta_0)^2;$
3. Check for wrong assignments and recalculate the structure.
4. Selcet 20 structures of least NOE violation (> 0.5 Å)
5. Criteria for good structures:
   a. No NOE violation
   b. RMSD < 0.5 Å
   c. No violation in dihedral angle (Inspect Ramanchandran diagram)(Atomic hindrance).

Overlay of 20 backbone Cα traces

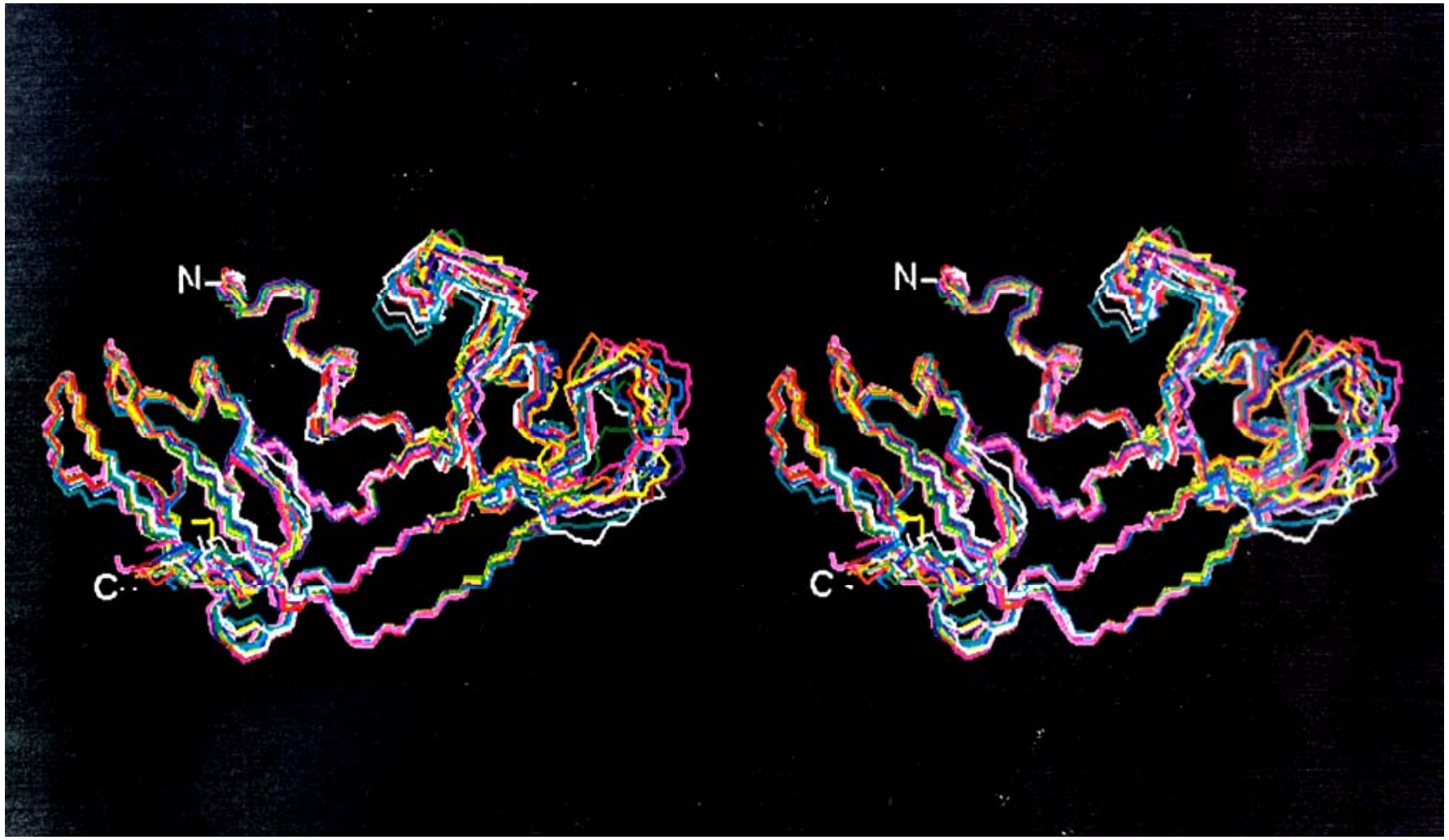
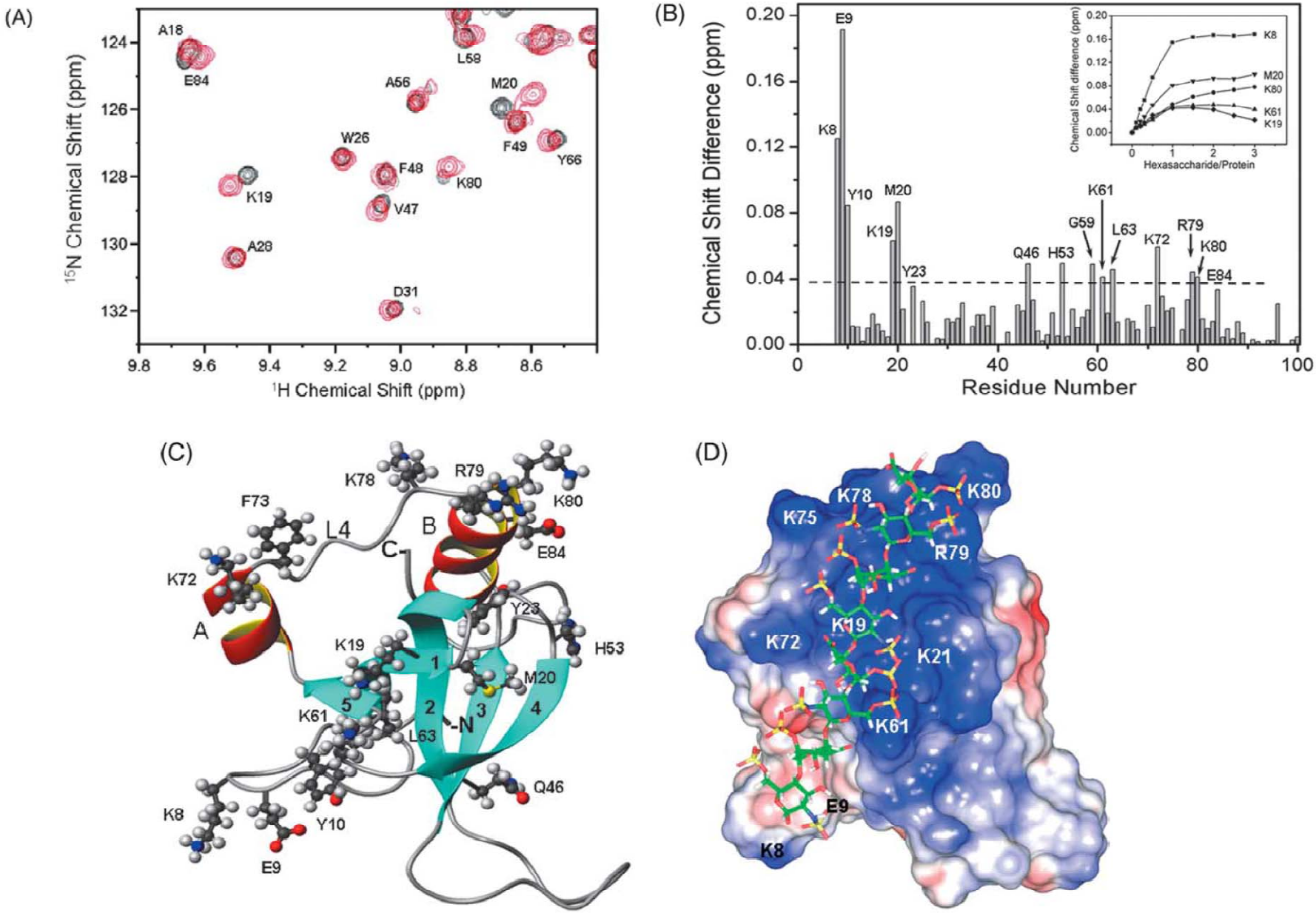


Table 1. Summary of structural constraints and structural statistics

NMR-derived restraints	
Upper inter-proton restraints	1245
Intra-residue	460
Sequential	435
Medium-range	146
Long-range	204
Hydrogen bond restraints	66 for 33 hydrogen bonds
Dihedral angle (ϕ, ψ, χ^1)	87, 81, 28
Total constraints	1507
Residual constraint violations[a]	
CYANA target function value (Å^2)	1.58 ± 0.31
NOE upper distance constraint violations	
Maximum (Å)	0.36 ± 0.20
Number >0.2 Å	3 ± 2
Dihedral angle constraint violations[b]	
Maximum (deg.)	4.86 ± 1.14
Number >5 (deg.)	1 ± 1
van der Waals violations	
Maximum (Å)	0.17 ± 0.02
Number >0.1 Å	0 ± 0
Average pairwise r.m.s. deviations (Å)[c]	
Backbone, N, C^α, C (10–93)	0.67 ± 0.12
Heavy atoms (10–93)	1.17 ± 0.18
Backbone, N, C^α, C (secondary region)[d]	0.39 ± 0.06
Heavy atoms (secondary region)	$0.98 + 0.12$
Ramachandran statistics	
Most favorable region (%)	73.1
Additional allowed region (%)	24.8
Generally allowed region (%)	2.0
Disallowed region (%)	0.1

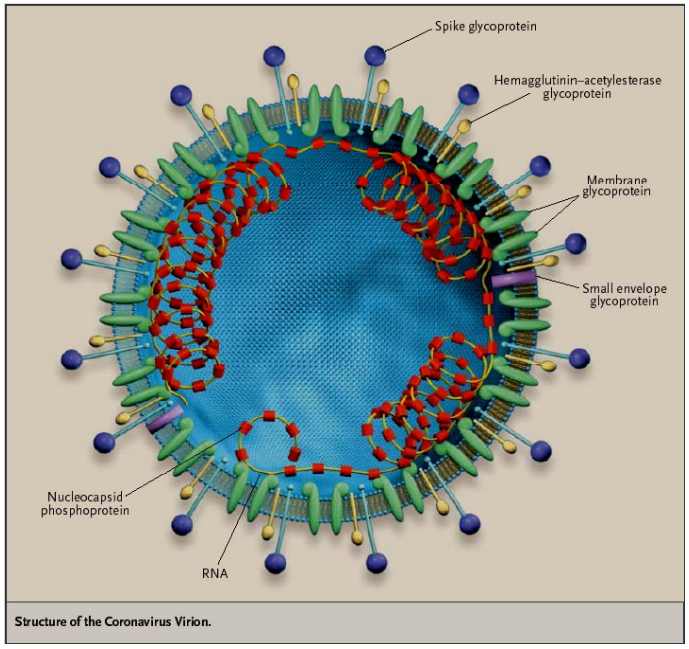
Mapping ligand binding site by chemical shift perturbation

(Heparin binding to HDGF)



Example 1

Dissecting the **Molecular Mechanism of the RNA Packaging of SARS Coronavirus Nucleocapsid**



Structure of the Coronavirus Virion.

Tai-huang Huang
IBMS, Academia Sinica

References

- Hsieh, P.K. et al (2005) "The Assembly of an RNA Packaging Signal Containing Virus-Like Particle of SARS-CoV is Nucleocapsid – dependent" J. Virol. **79**, 13848-13855.
- Chang, C.K. et al (2005) "The dimer interface of the SARS coronavirus nucleocapsid protein adapts a porcine respiratory and reproductive syndrome virus-like structure" FEBS Lett. **579**, 5663-5668
- Chang, C.K. et al (2006) "Modular organization of SARS coronavirus nucleocapsid protein". J. Biomed. Sci. 13, 59-72.
- Chen, C.Y. et al. (2007) "Structural insight into the helical packing from the crystal structure of SARS coronavirus nucleocapsid protein dimerization domain.
- Takeda, M et al. (2008) "Solution Structure of the C-terminal Dimerization Domain of SARS Coronavirus Nucleocapsid Protein Solved by the SAIL-NMR Method" J. Mol. Biol. **380**, 608-622.

Severe Acute Respiratory Syndrome (SARS)
– The first pandemic thread of the 21st century

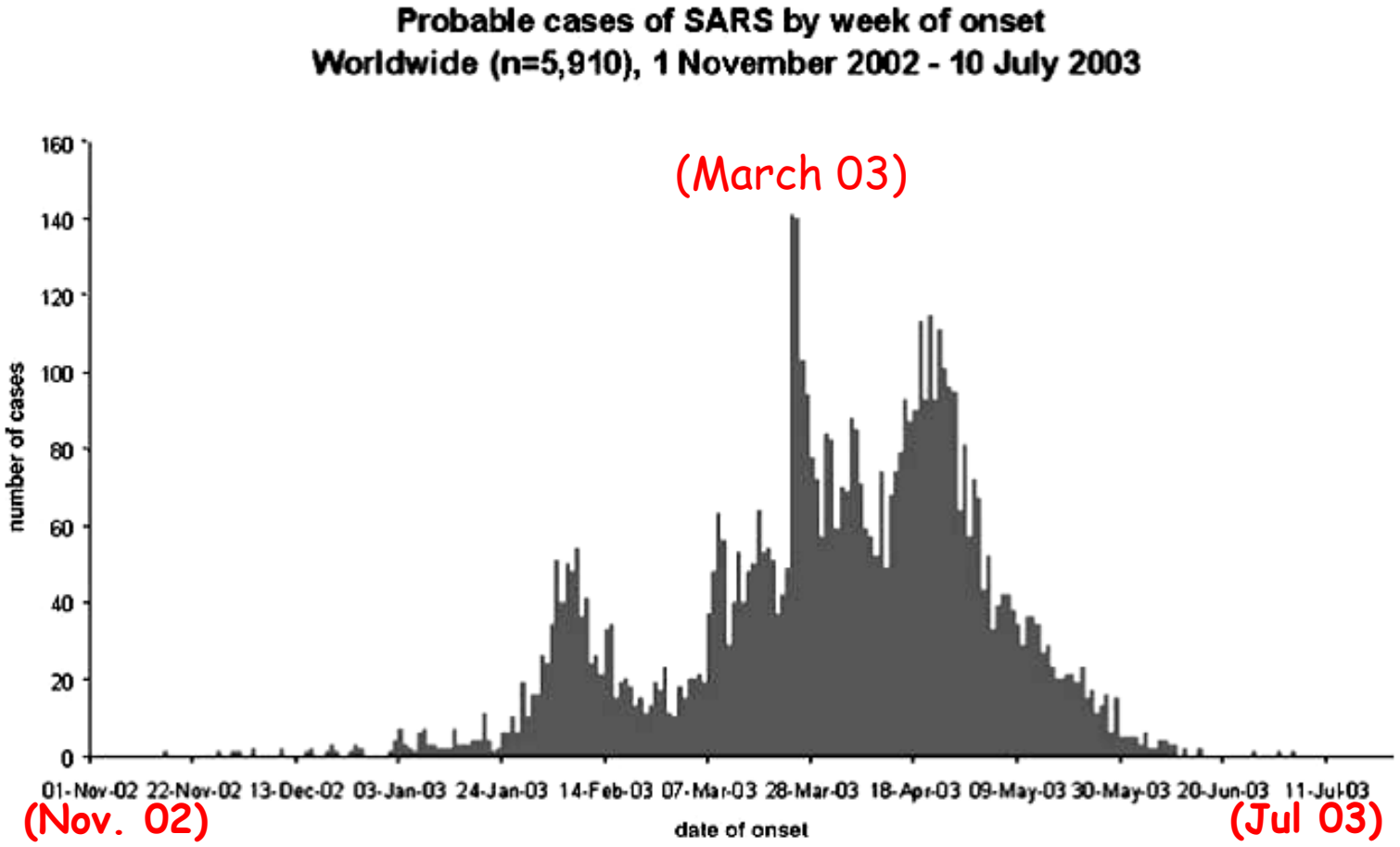
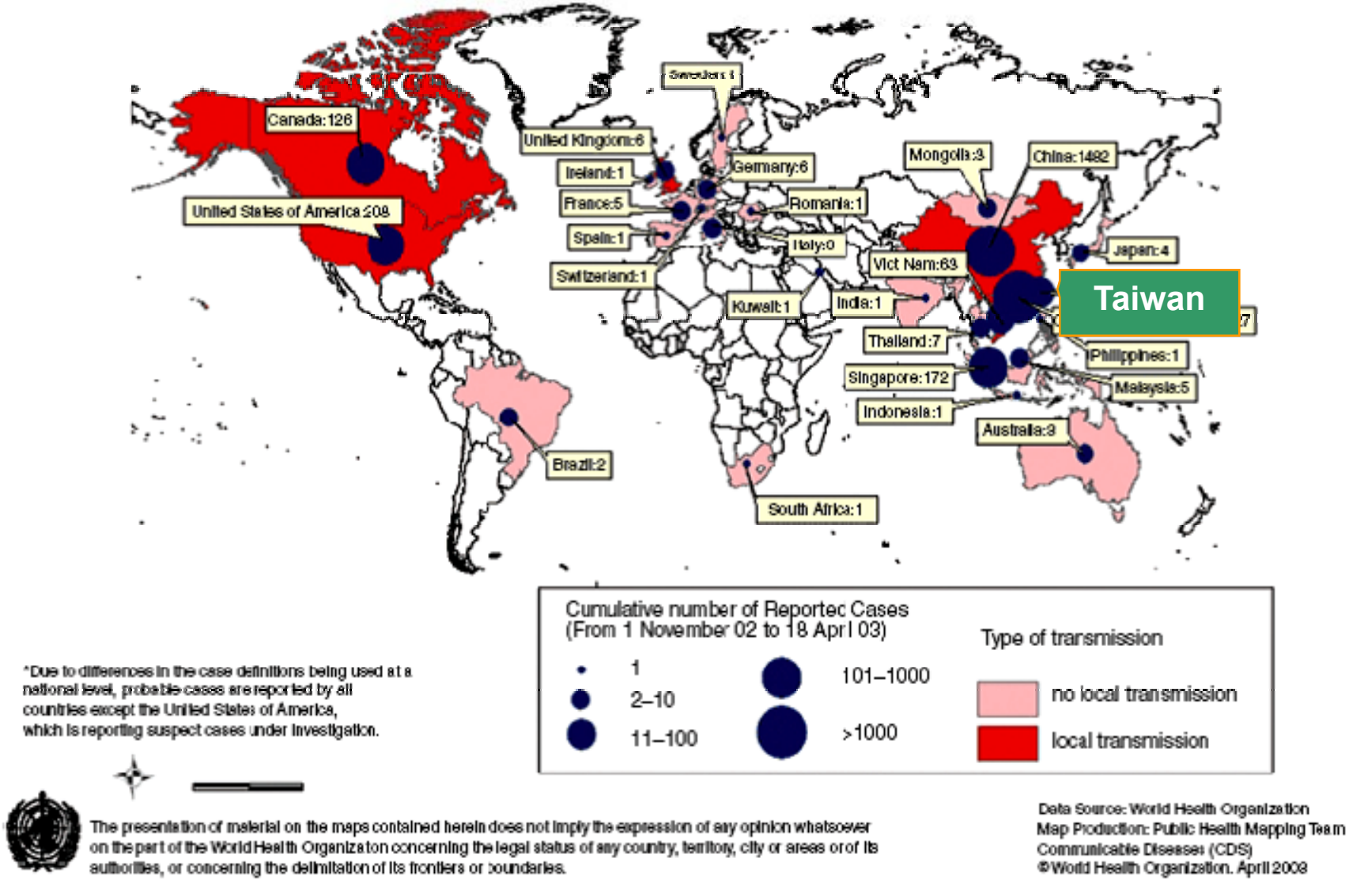


Figure 1 Pandemic curve of probable cases of severe acute respiratory syndrome. This graph does not include 2,527 probable cases of SARS (2,521 from Beijing, China), for whom no dates of onset are currently available. Source: http://www.who.int/csr/sars/epicurve/epiindex/en/index1.html.

Skowronski *et. al.* (2005) Ann. Rev. Microbiol. 56, 357-381)

Globally, 8,098 people were infected from 29 countries, 774 casulties.

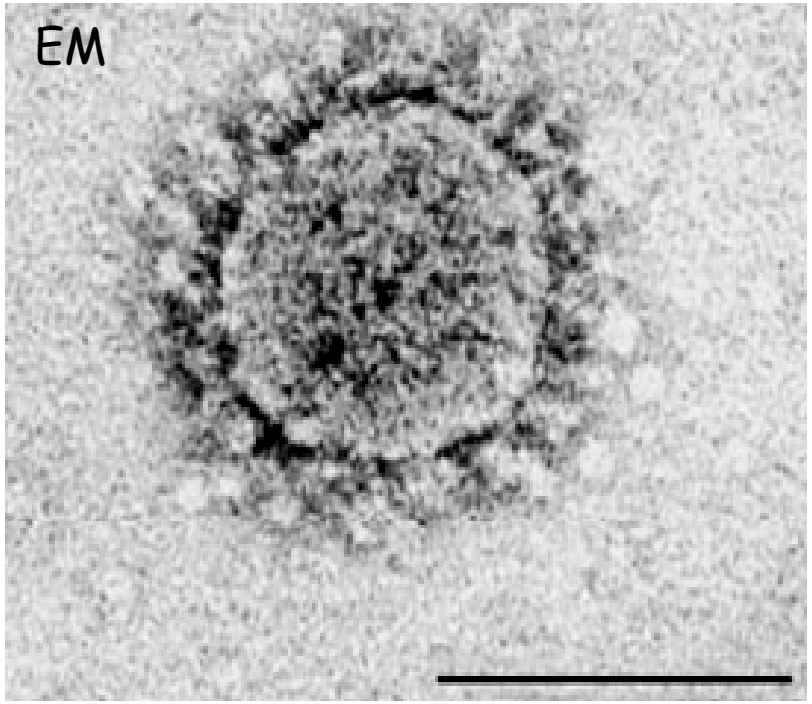
Economic impact due to travel and investment ~ US$30-140 billions)

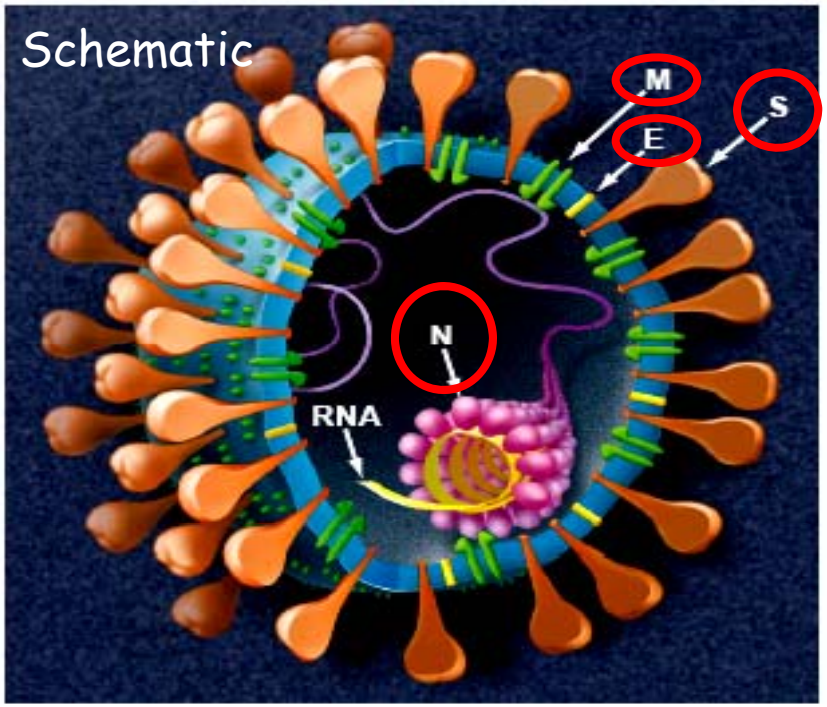


Scandinavian Journal of Immunology 58, 277–284

Causative agent – SARS Coronavirus

1. A single stranded plus-sense enveloped RNA virus.
2. Genome of 29,751 nt, containing 14 ORF encoding 28 proteins





Four Structural proteins:

S: Spike protein (1255 a.a.);
E: Envelope protein (76 a.a.)
M: Membrane protein (221)
N: Nucleocapsid protein (422 a.a.)

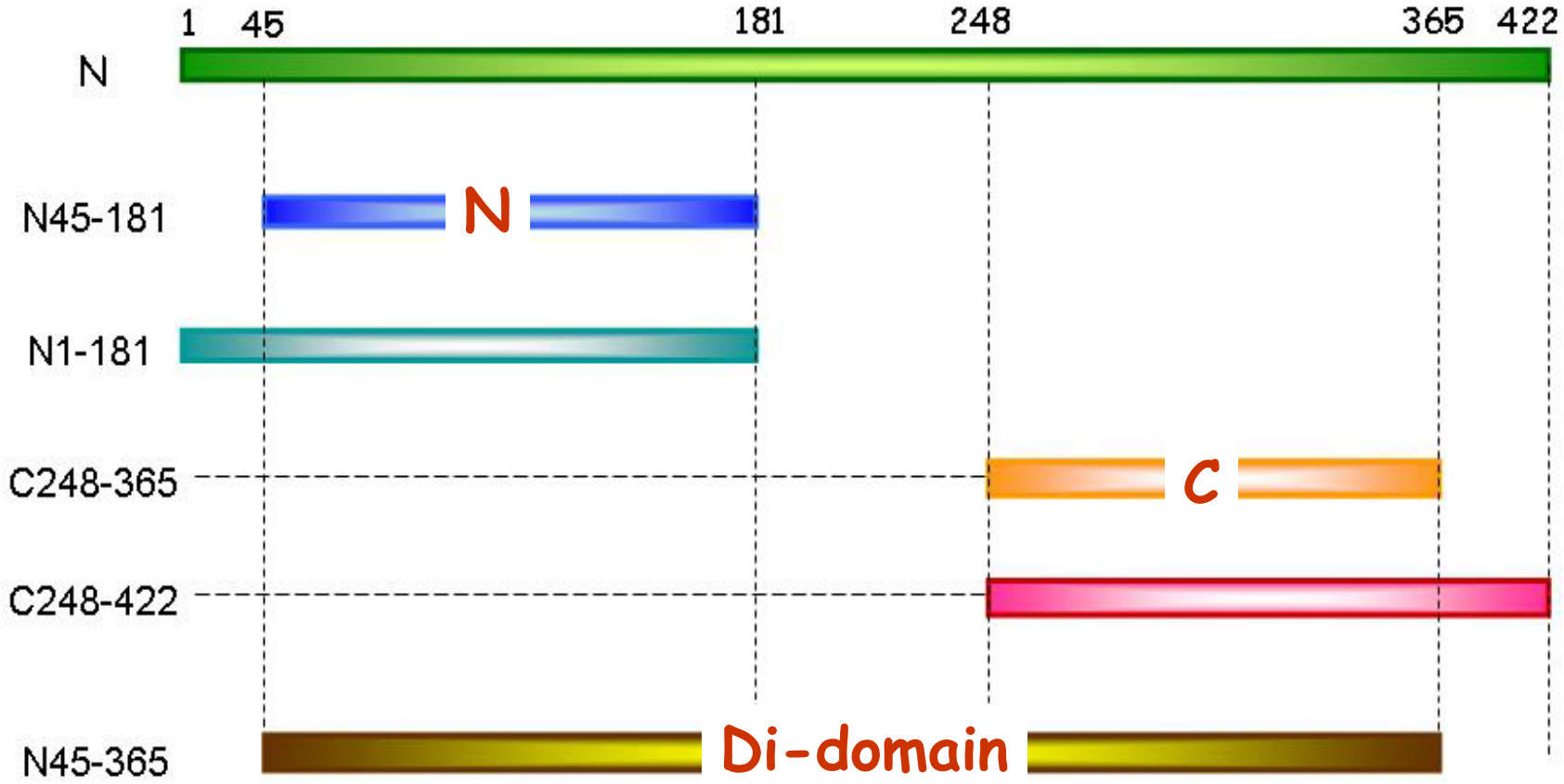
SARS nucleocapsid protein N (422 a.a.)

- The most abundant viral protein and a major antigenic determinant ➔ **Target for detection and vaccine developments.**
- Bind to RNA to form a helical RNP and it interacts with M protein in stabilizing the nucleocapsid
  ➔ **Important in virion assembly, packaging and release.**
- Interact with various host protein systems and implicated in several functions such as replication and apoptosis etc:
  - Interacts with AP-1 signal transduction pathway ?
  - Interacts with Smad3 and Modulatest transforming Growth Factor- Signalin (JBC-2008)
  - Inhibits Cell Cytokinesis and Proliferation by Interacting with Translation Elongation Factor 1α (JV-2008)

Dissecting the domain structure of the SARS CoV N protein

1. Spectrum of full length protein was terrible
   ➔ Divide and Conquer

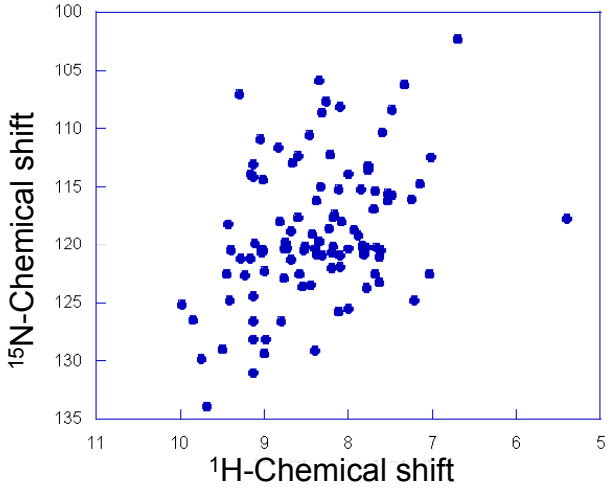
2. Total of 28 clones have been constructed and expressed.

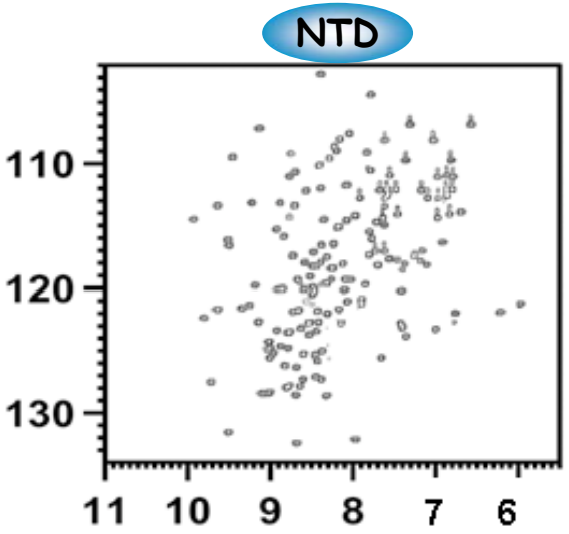


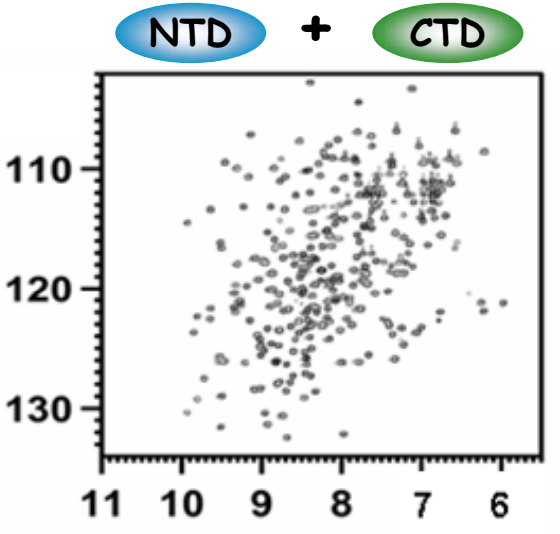
Characterization of protein disorder by NMR

^{15}N-Heteronuclear Single Quantum Correlated Spectroscopy (^{15}N-HSQC)

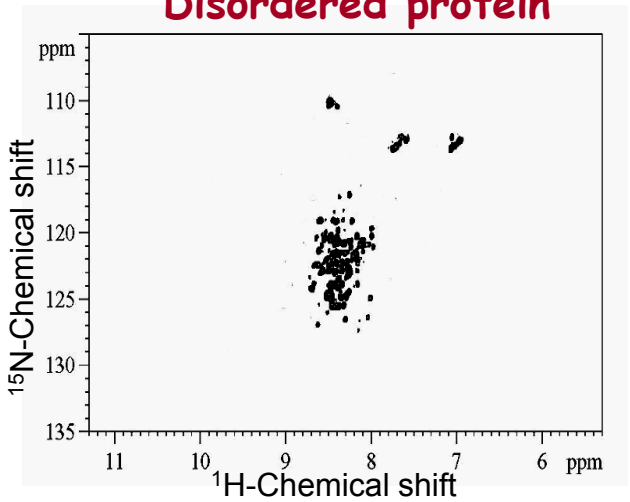
Folded protein

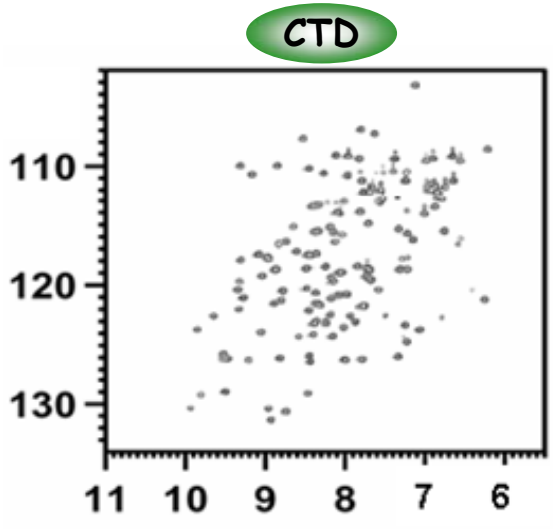




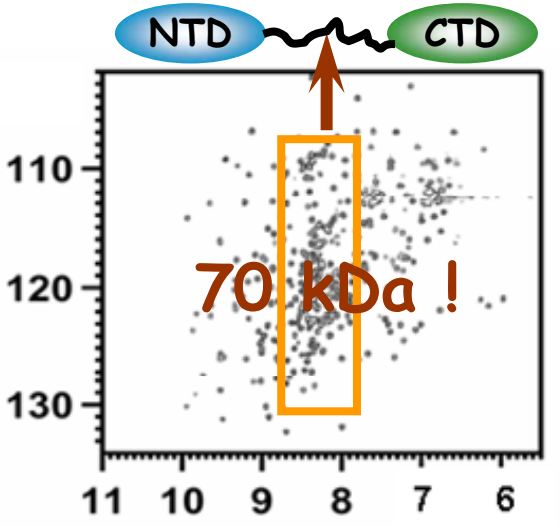


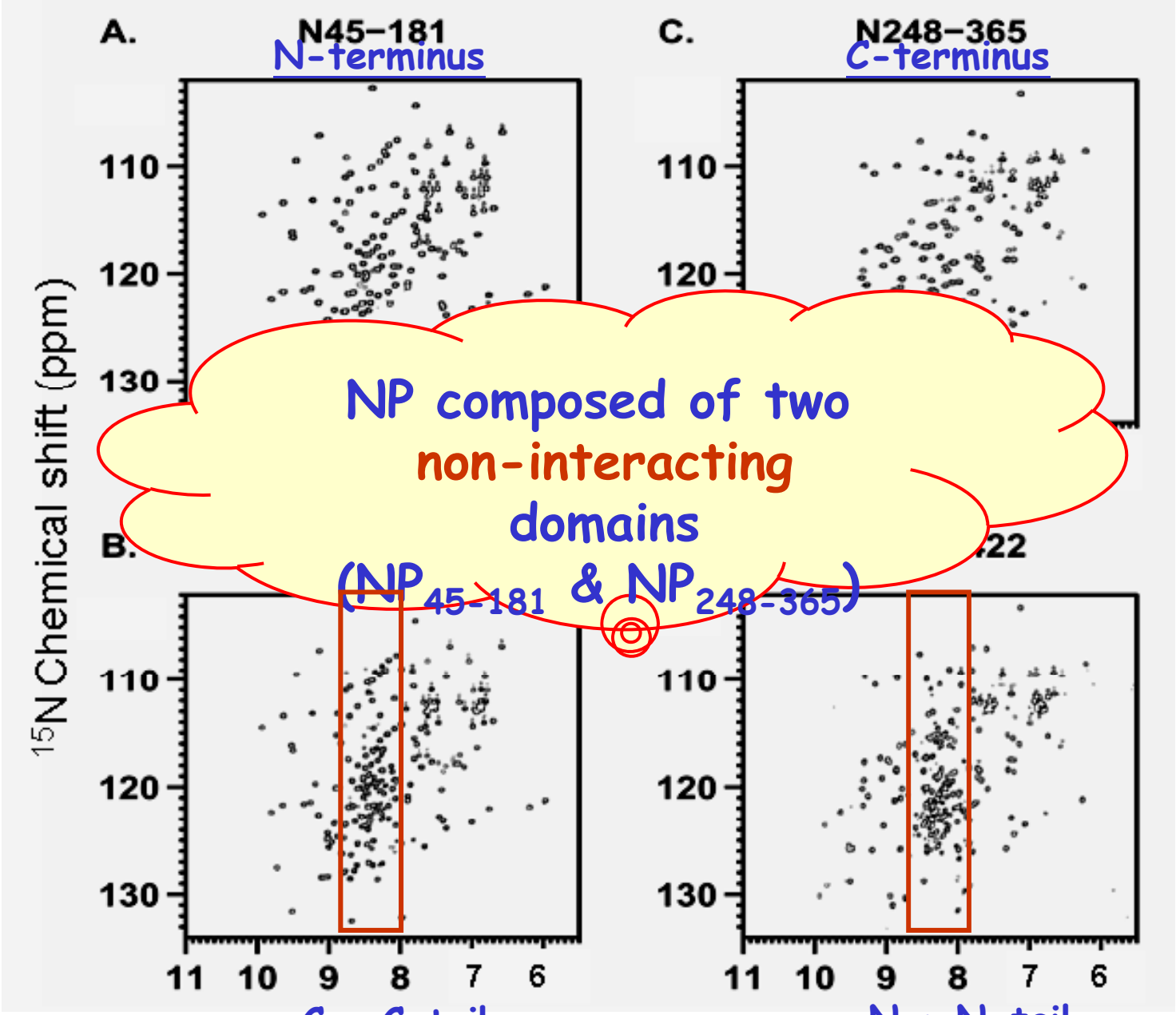
Disordered protein





Disordered linker

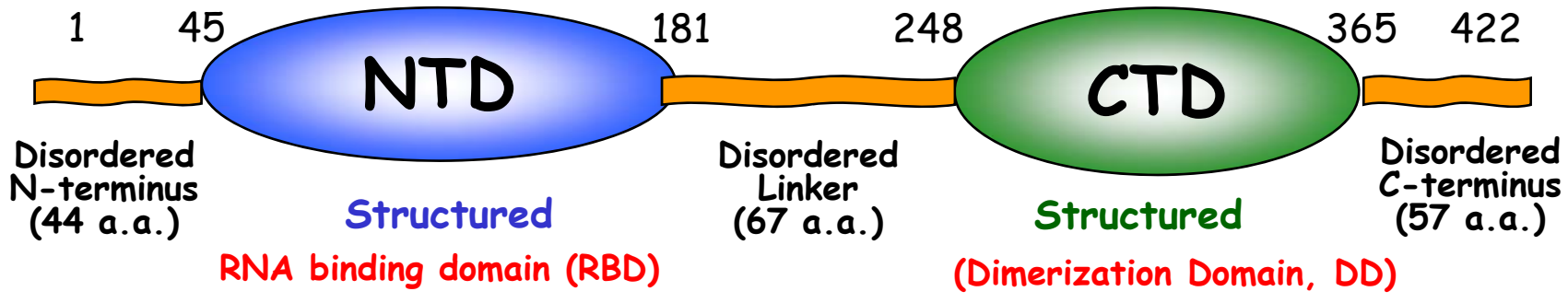


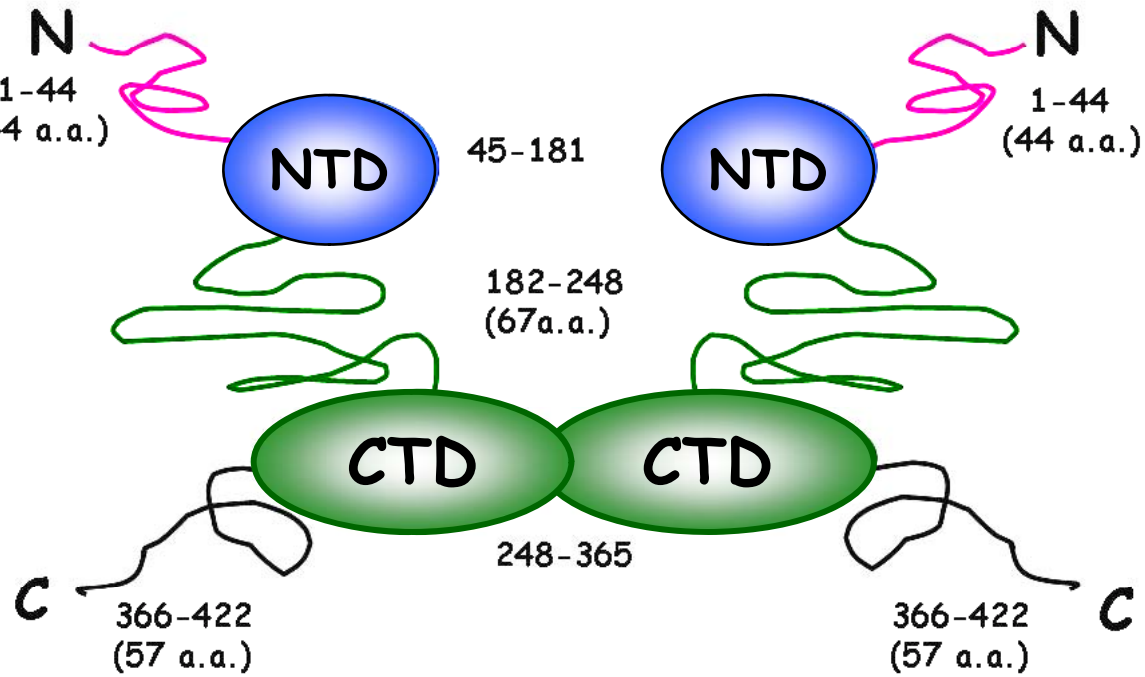

A.
N45–181
N-terminus
C.
N248–365
C-terminus
B.
C + C-tail
N + N-tail
^{15}N Chemical shift (ppm)
110
120
130
11
10
9
8
7
6
NP composed of two
non-interacting
domains
(NP_{45-181} & $NP_{248-365}$)

Evidences that $N_{248-365}$ forms a dimer

1. Size exclusion chromatography.
2. NMR relaxation (short T_2).
3. Chemical cross-linking
4. Ubiquitin-OGD chimeric protein mixing expt.
5. Light scattering.
6. Analytical ultracentrifugation.

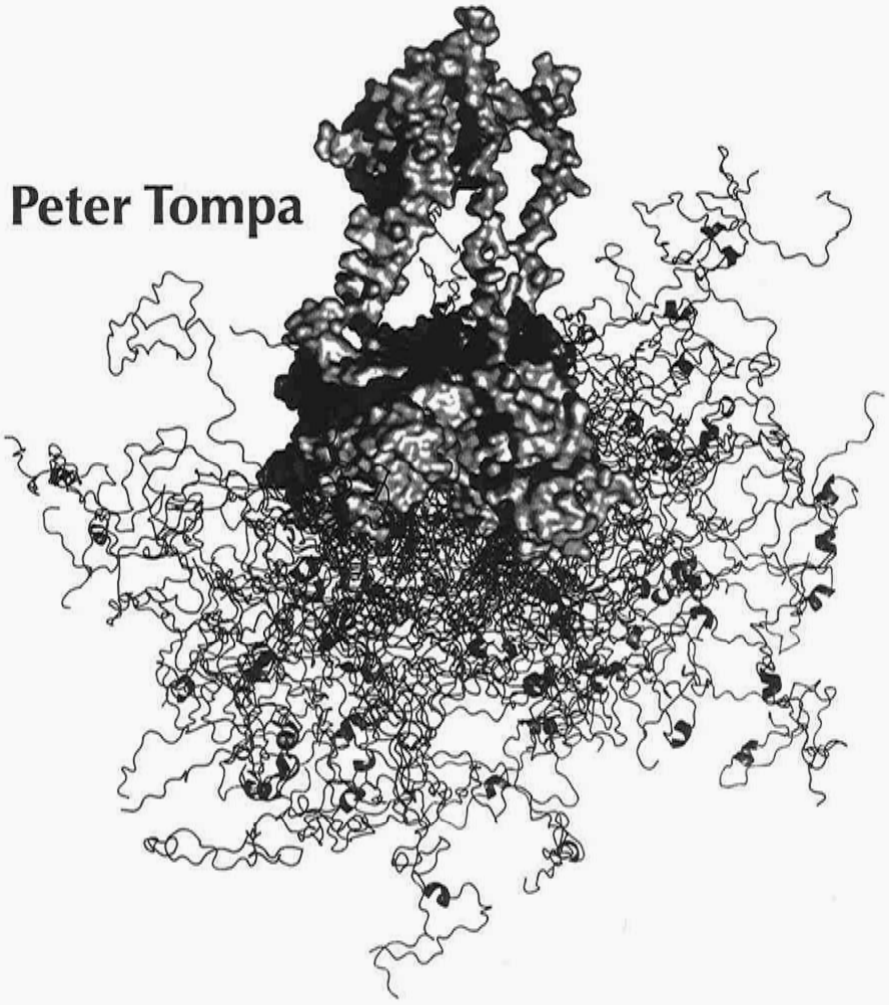
Domain architecture of SARS-CoV NP





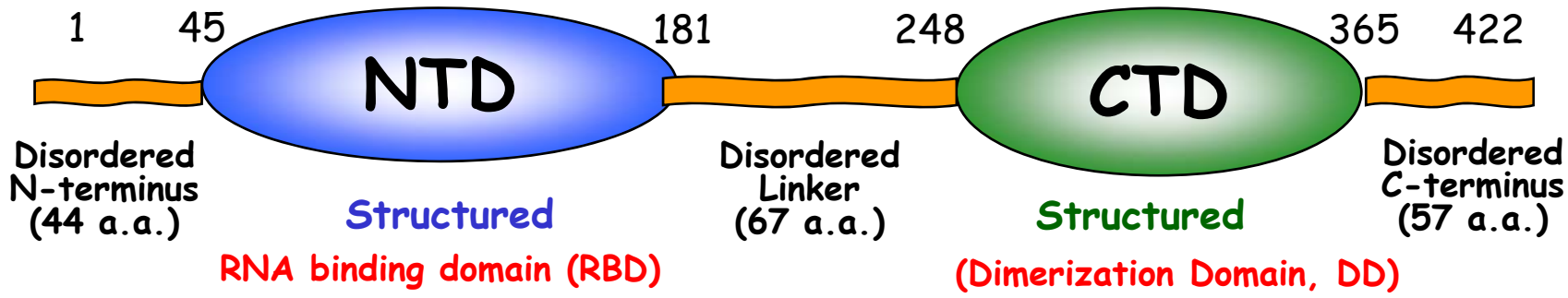
Structure and Function
of Intrinsically Disordered Proteins

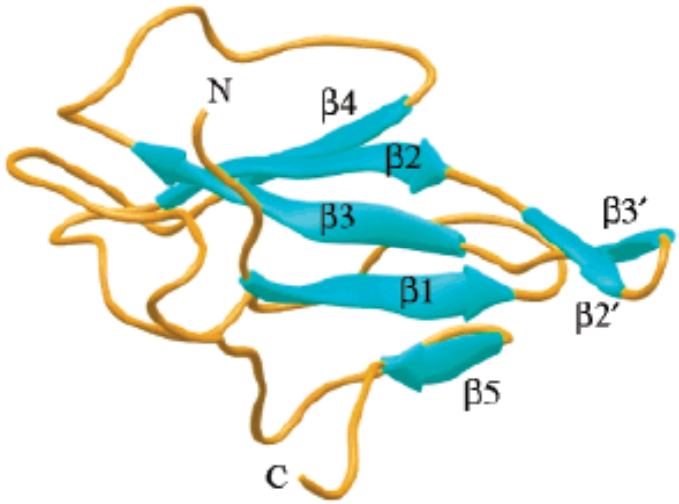
Peter Tompa



CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Tertiary structure

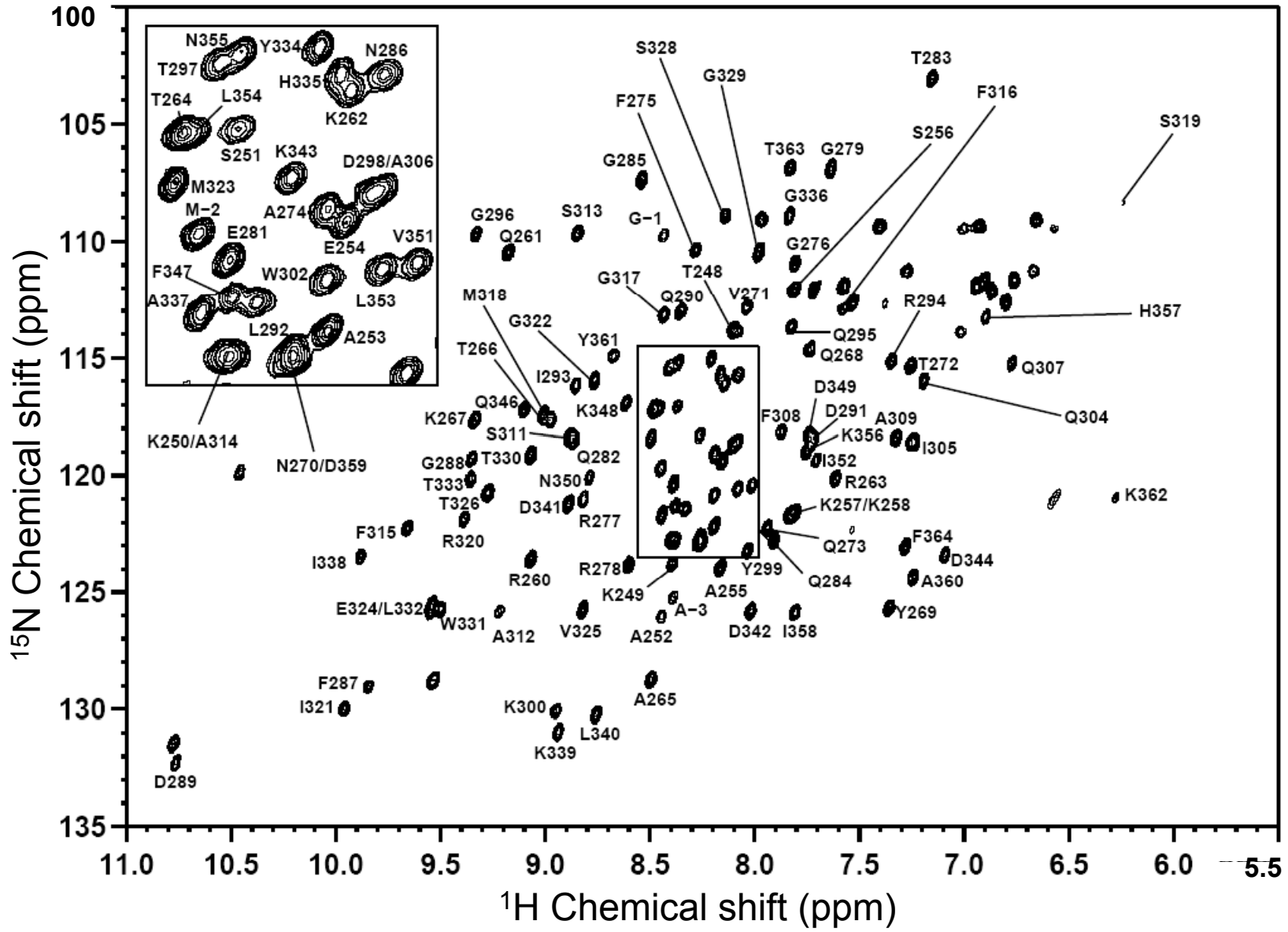




?

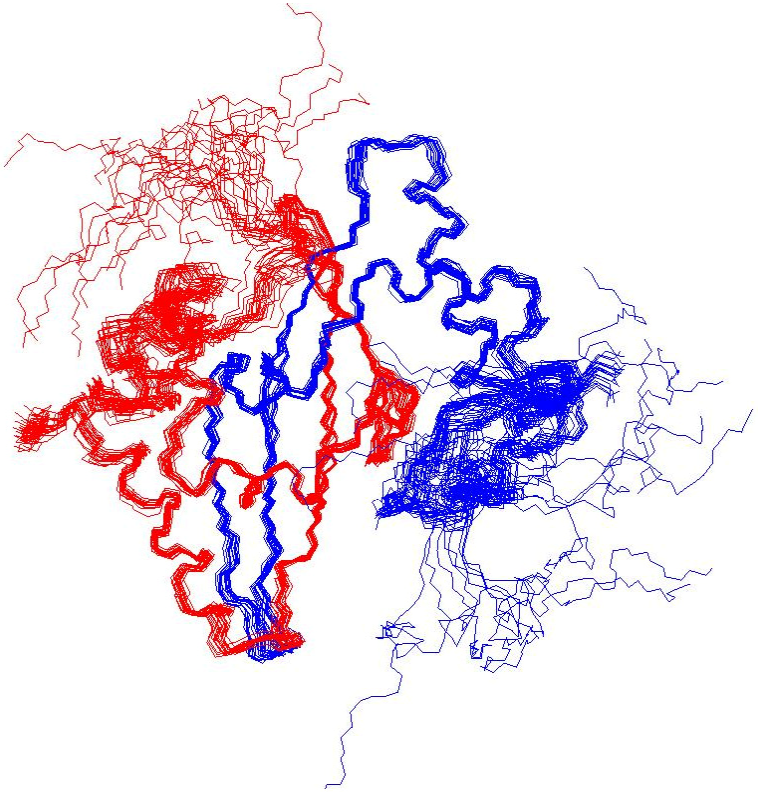
(Solved by Abbott group, May 2004)

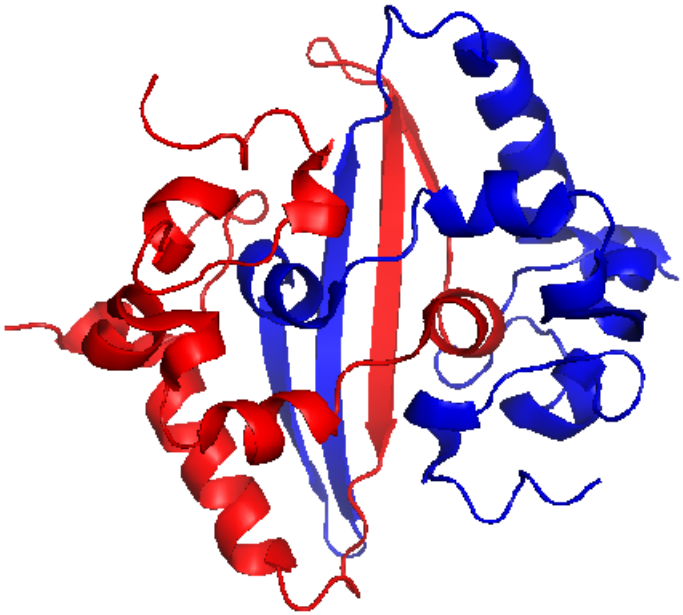
Structural determination – NMR resonance assignments



NMR structure of SARS-CoV NP CTD

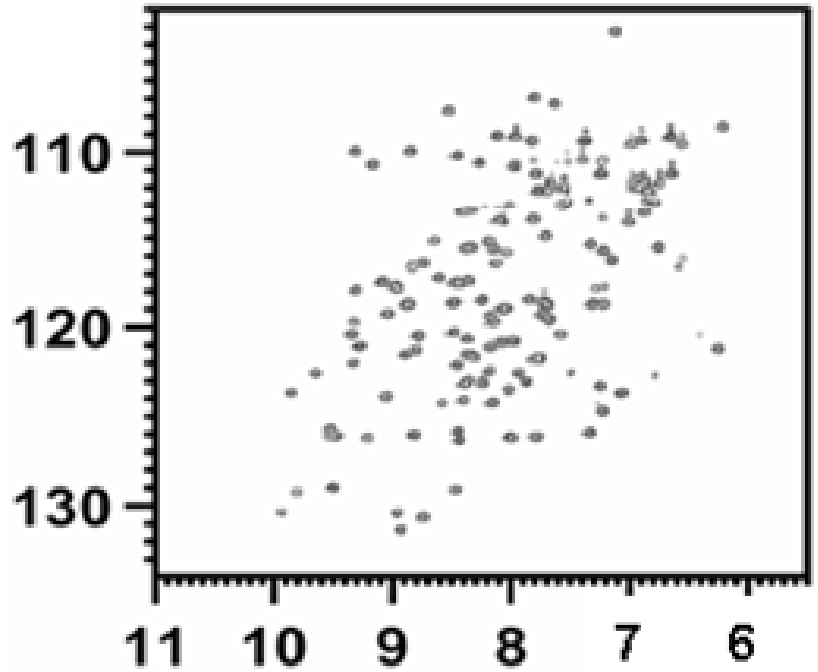
(28 kDa homo-dimer solved by SAIL method)

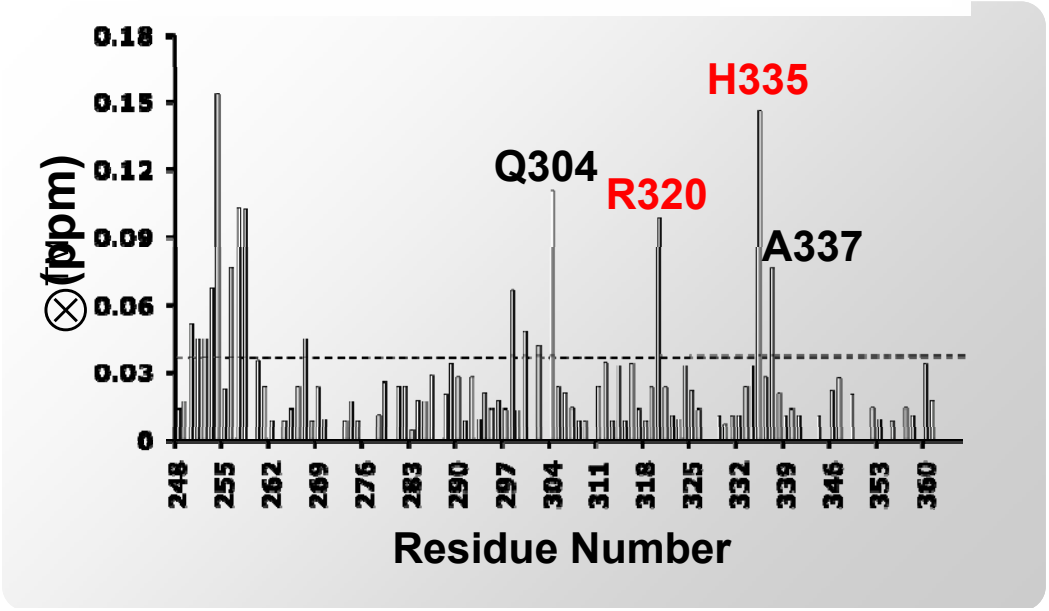


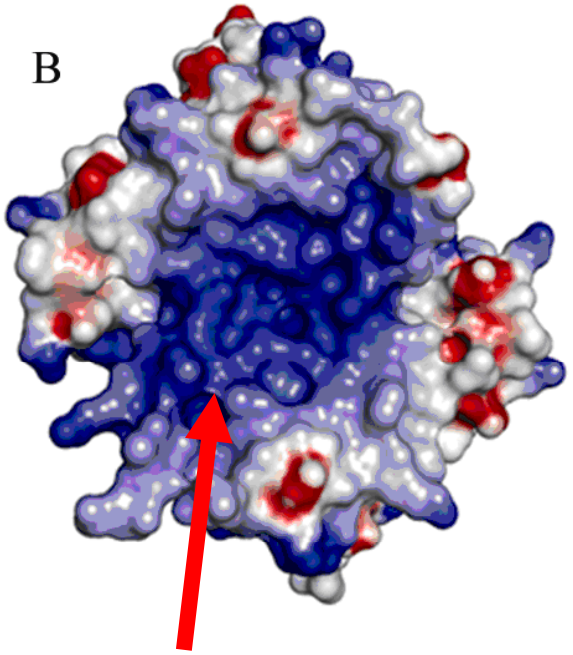


(Collaboration with M. Kainosho of Kyoto U)

Identification of RNA binding site in CTD





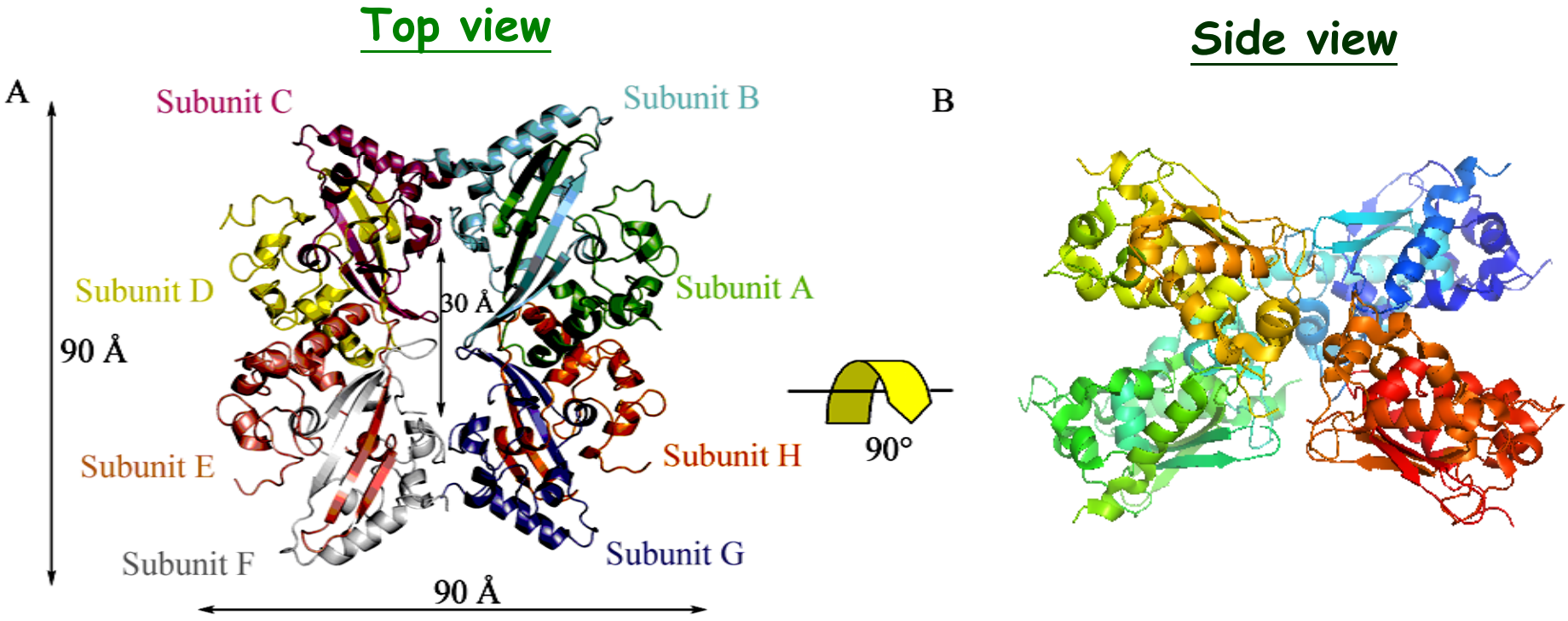


Primary RNA binding site.

IV. RNP Packaging

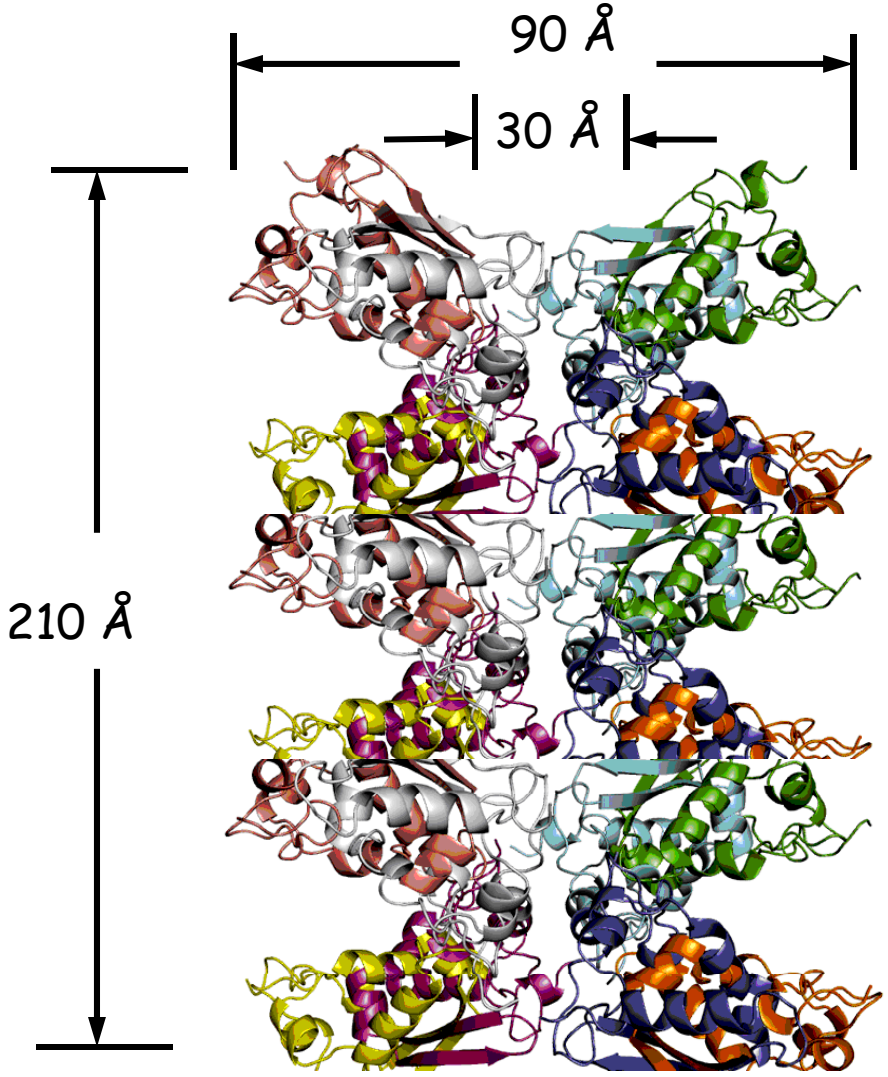
X-ray crystallography

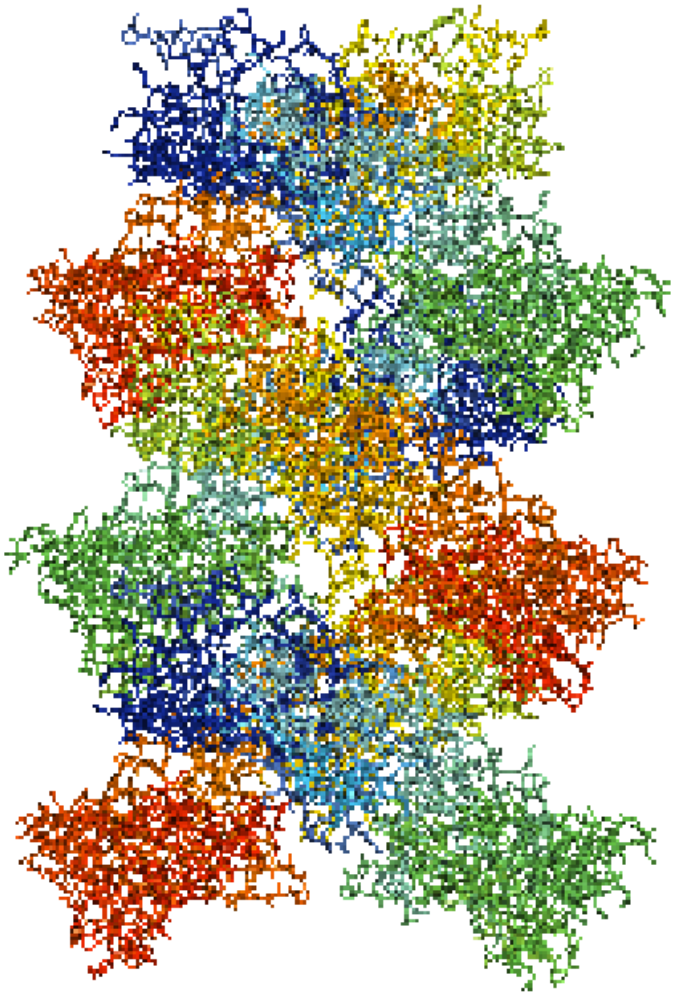
- CTD packs as an octamer in an unit cell



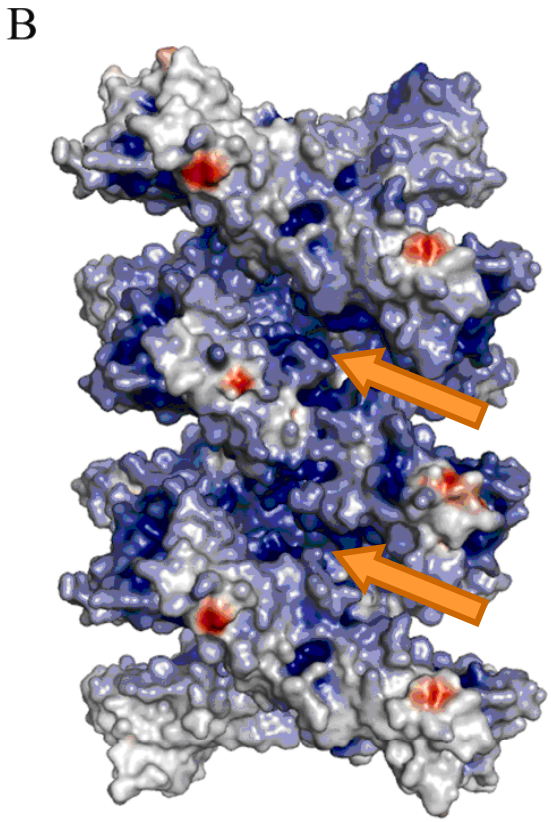
Crystal packing

➔ Stacking of 3 octamers forms a complete turn of a left-handed twin helix.





Surface Charge Potential

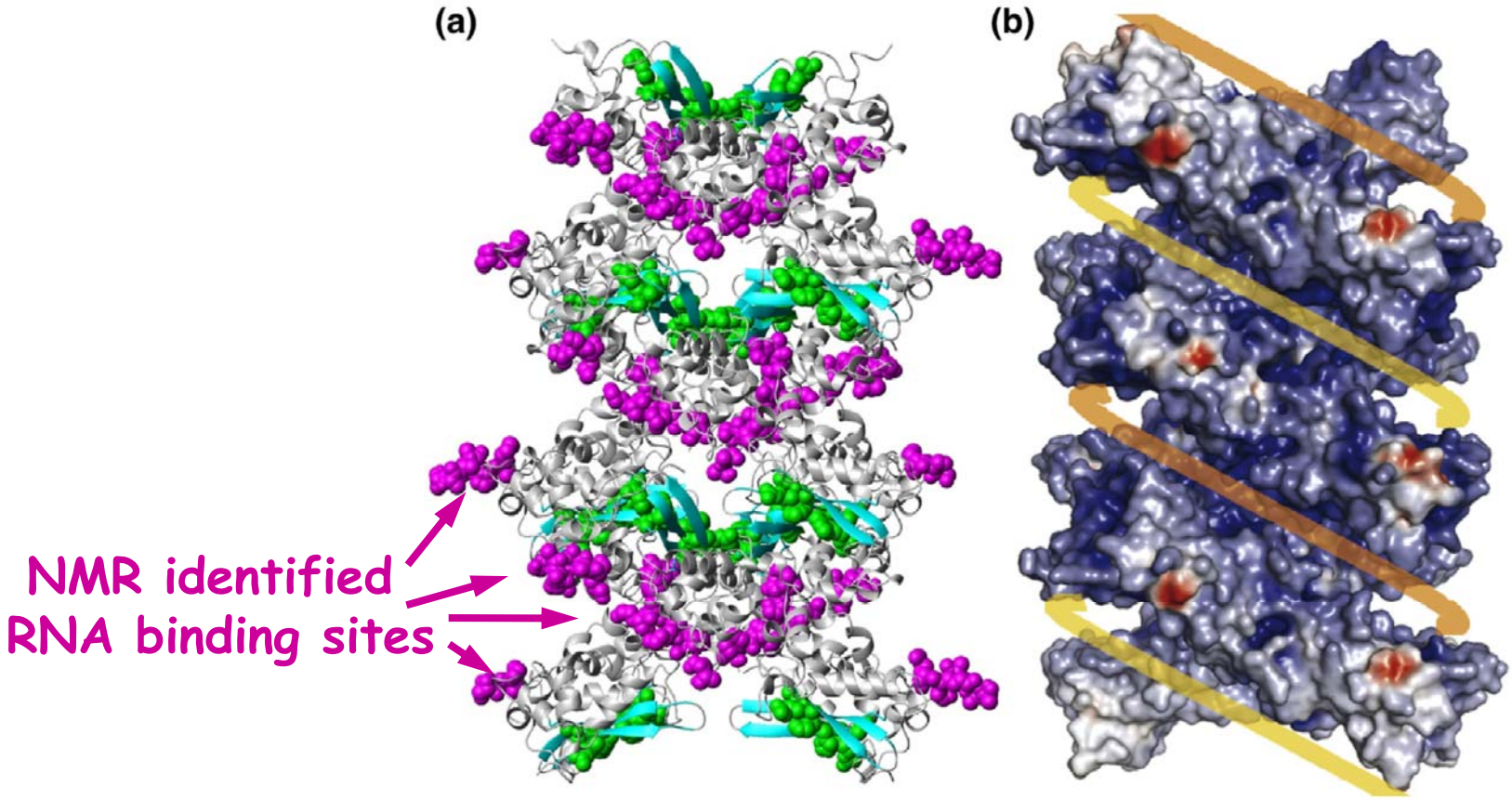


Left-handed twin helix

"+" charged grooves are presumed to be the nucleotide binding sites

➔ We propose that RNA binds to the Left-handed twin helix grooves non-specifically to form helical RNP.

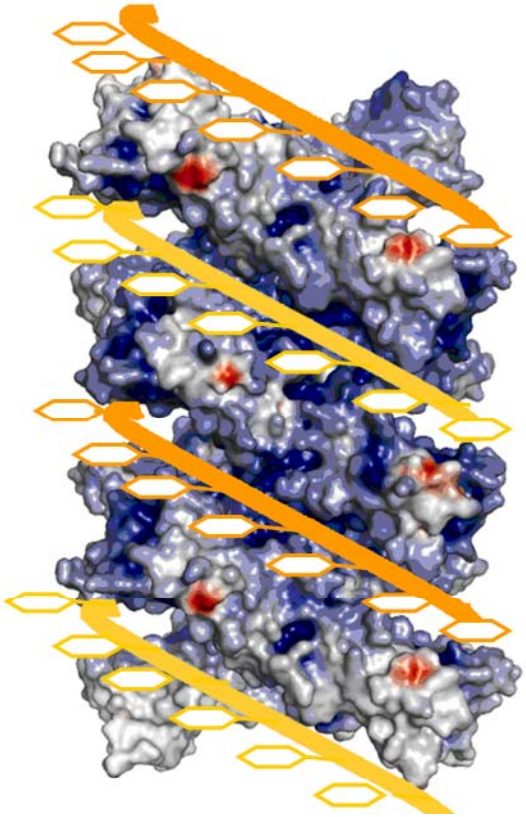
Confirmed by NMR

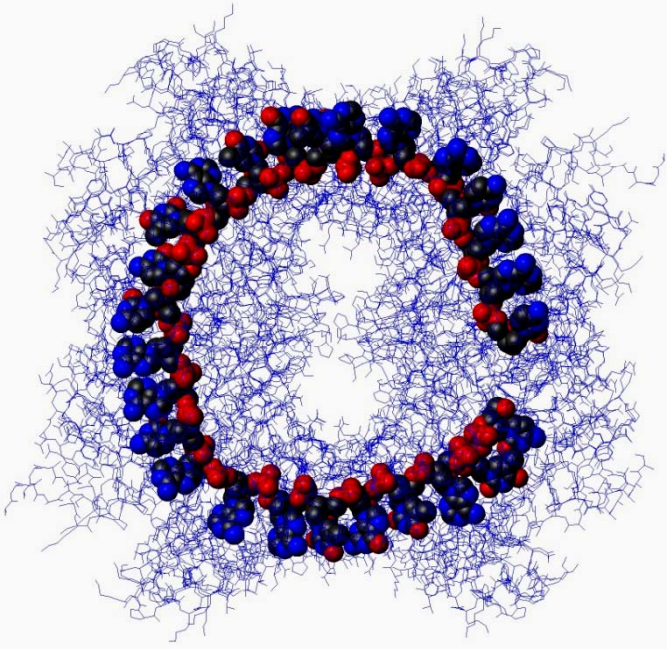


Proposed model of the NP/RNA complex

➔ RNA wrap around the CTD to form a left-handed twin-helix.

➔ Bases are facing outward.





Role of NTD – Cap the exposed bases

- ➔ CTD forms the helical core.
- ➔ NTD covers the exterior and interacts with the bases.