Supporting Information

Predominant Role of Water in Native Collagen Assembly inside Bone Matrix

Ratan Kumar Rai[1], *Chandan Singh*[1,2] *and NeerajSinha*[1]*

[1]Centre of Biomedical Research
SGPGIMS Campus, Raibarelly Road
Lucknow – 226014 INDIA

[2]School of Biotechnology
Banaras Hindu University
Varanasi – 221005 INDIA

* Corresponding author

e-mail: neeraj.sinha@cbmr.res.in , neerajcbmr@gmail.com

Fax: +91-522-2668215

Methods

NMR experimental parameters

All ssNMR spectra were recorded on 600 MHz NMR spectrometer (Avance III, BrukerBiospin, Switzerland) operating at 600.154 MHz for ^{1}H, and 150.154 MHz for ^{13}C frequencies with Bruker's 3.2 mm Efree probe. The MAS spinning speed was controlled by Bruker's MAS pneumatic unit within accuracy of ± 2Hz. Pulse sequencesare given in Figure S2 andexperimental parameters aresummarized in Table S1.

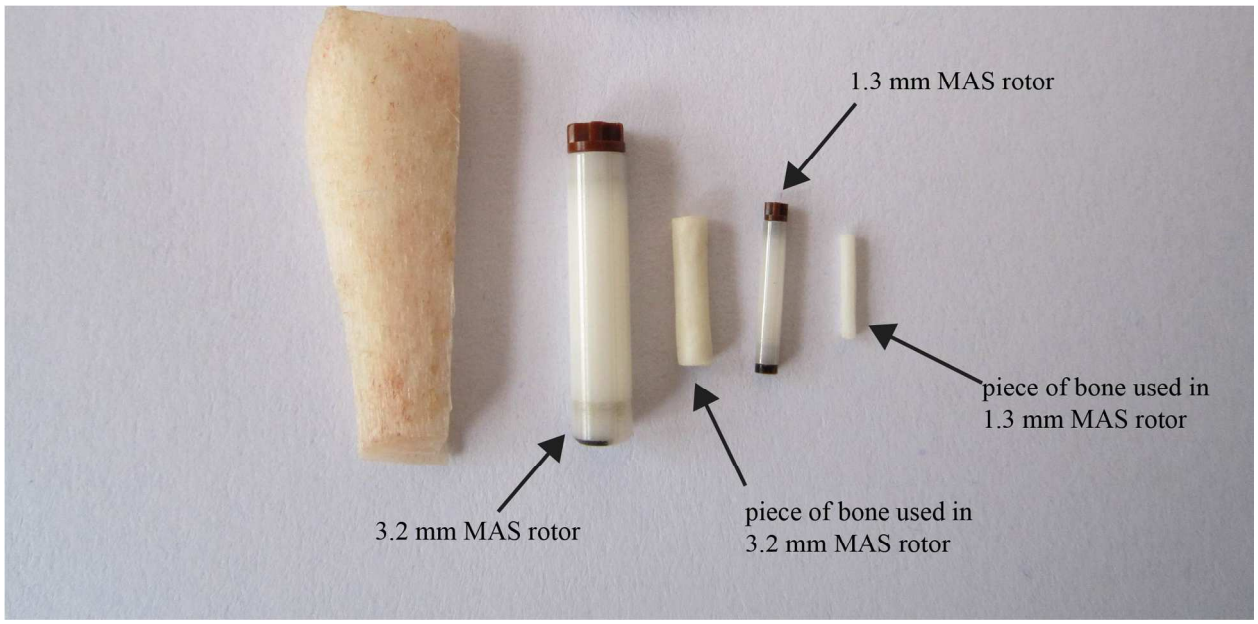

Figure S1. Picture of intact bone,different pieces of bone and MAS rotors.

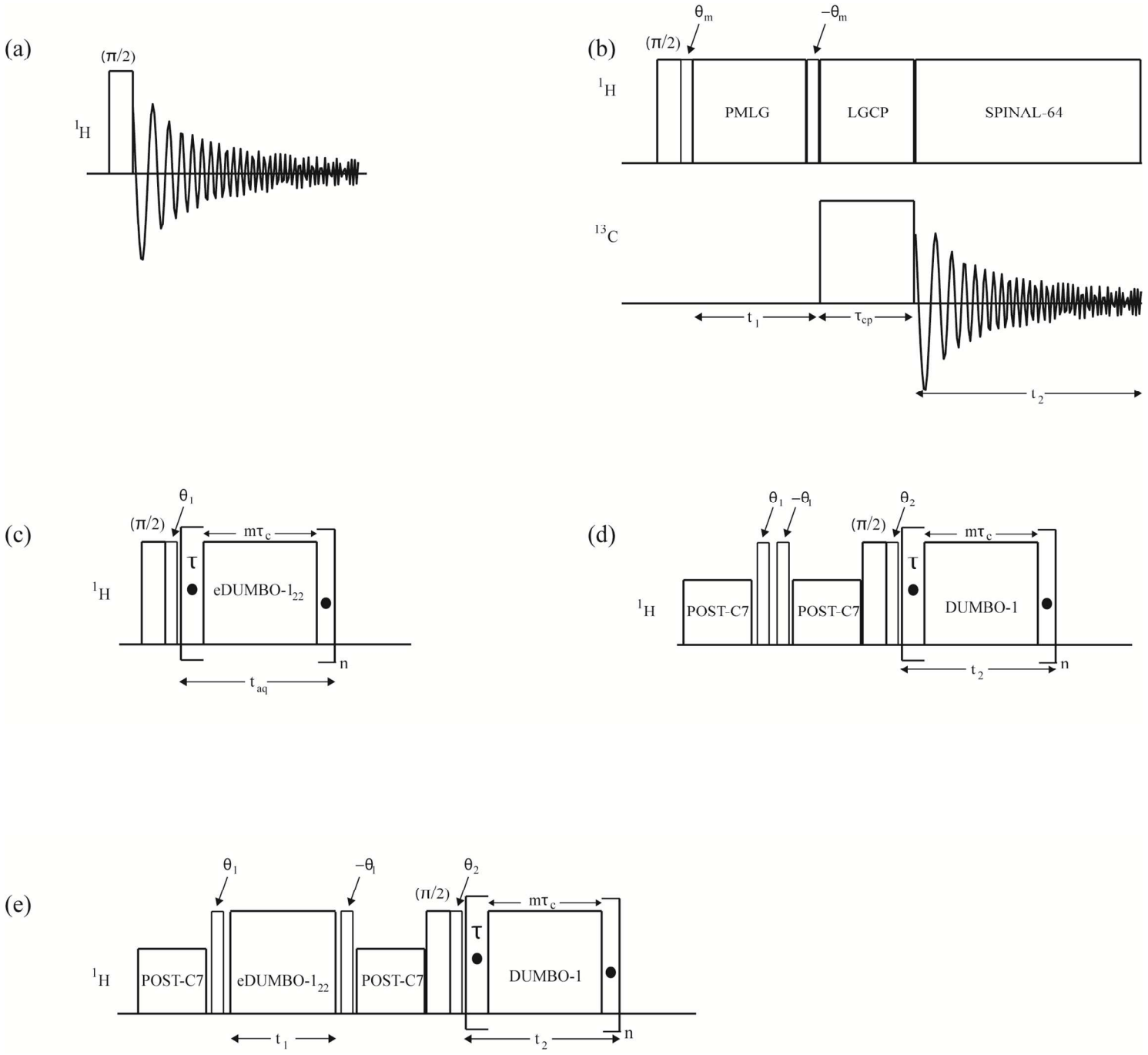

Figure S2. Pulse sequences used for different ssNMR experiments.(a) One pulse ^{1}H NMR (b) ^{1}H/^{13}C HETCOR (c) ^{1}H 1D DUMBO (d) ^{1}H DQ-SQ 1D and (e)^{1}H - ^{1}H DQ-SQ CRAMPS pulse sequences.

Table S1: NMR spectroscopy experimental parameters:

S.No	NMR experiment	Probe/spinning speed	Experimental parameter
1	One pulse ^{1}H 1D spectrum (Figure S2(a))	3.2mm Efree / 10 kHz	The ^{1}H $\pi/2$ pulse for one pulse ^{1}H NMR experiment was 2.4 μs. It was recorded with 1k data point, acquisition time of 10.2 ms for each bone sample. We recorded 32 scans with recycle delay of 3seconds. (Figure 1)
2	^{1}H/^{13}C HETCOR (Figure S2 (b))	3.2mm Efree / 10 kHz	For ^{1}H−^{13}C LGCP-HETCOR experiments, the LG-CP contact time was 70 μs and the maximum t_1 evolution time was 1.6 ms[1]. The effective field during ^{1}H homonuclear decoupling period (PMLG) was 80 kHz and high power ^{1}H decoupling (100 kHz) was applied during t_2 period[1-3]. A total of 1k transients per increment and recycle delay of 3.5 seconds. Contact time during LG-CP was 70 μs to ensure one bond correlation. The spectra were zero filled, linearly predicated up to 32 points and sine bell apodizationwasused in both dimension prior to Fourier transformation. The spectra were referenced by internal reference to alanine chemical shift. Small quantity of powder alanine was filled inside rotor along with bone sample while recording NMR spectrum (Figure 2 and Figure S3.)
3	1D ^{1}H DUMBO (Figure S2(c))	3.2mm Efree / 12.5 kHz	The ^{1}H $\pi/2$ pulse was 2.4 μs. The RF field strength used for all pulses, including DUMBO decoupling shape pulse was 100kHz.(Figure 2). We recorded 128 scans with recycle delay of 1seconds.
4	^{1}H DQ-SQ 1D DUMBO (Figure S2(d))	3.2mm Efree / 12.5 kHz	The ^{1}H $\pi/2$ pulse was 2.4 μs. The RF field strength used for all pulses, including DUMBO decoupling shape pulse was 100kHz.We have used POST-C7 Block for the excitation and reconversion of double quantum[4]. This sequence was used to record double quantum filtered single quantum (DQ-SQ 1D). We recorded 128 scans with recycle delay of 3 seconds. Figure 1(b)
5	^{1}H DQ-SQ CRAMPS (Figure S2(e))	3.2mm Efree/12.5 kHz	The ^{1}H $\pi/2$ pulse was 2.4 μs. The RF field strength used for all pulses, including the DUMBO decoupling shape pulse was 100kHz. The θ_1 pulse was calibrated carefully to 1.3 μs. We have used POST-C7 Block for the

			excitation and reconversion of double quantum[4]. The excitation and reconversion periods were 68.6 μs (corresponding to 3 basic POST-C7 elements). The proton RF field was set to 87 kHz during the POST-C7 blocks. DUMBO-1 blocks were 96 ms long (corresponding to three basic DUMBO-1 cycles).The t_1 increment was set to 96 ms (corresponding to three basic e-DUMBO-1_{22} cycles)[5,6]. The scaling factors for eDUMBO-1_{22} (during t_1) and DUMBO-1 (during t_2) were determined experimentally as 0.54 and 0.50, respectively. A total of 128 t_1 points were recorded for each experiment. For each of t_1 slices, 128 scans were recorded with a recycle delay of 3 seconds[7]. The data were zero filled and linearly predicted up to 32 points prior to Fourier transformation. (Figure 3) Decoupling sequence were optimized following protocol mentioned in literature[8]
6	WISE	3.2mm Efree/4 kHz	For $^{1}H/^{13}C$ wide line separation (WISE) experiments 50 us contact time was used with recycle delay of 2 seconds[9]. A total of 24 sampling point at 10 μs increment along t_1 were obtained. SPINAL-64 decoupling (100 kHz ^{1}H r.f. field) used during acquisition.

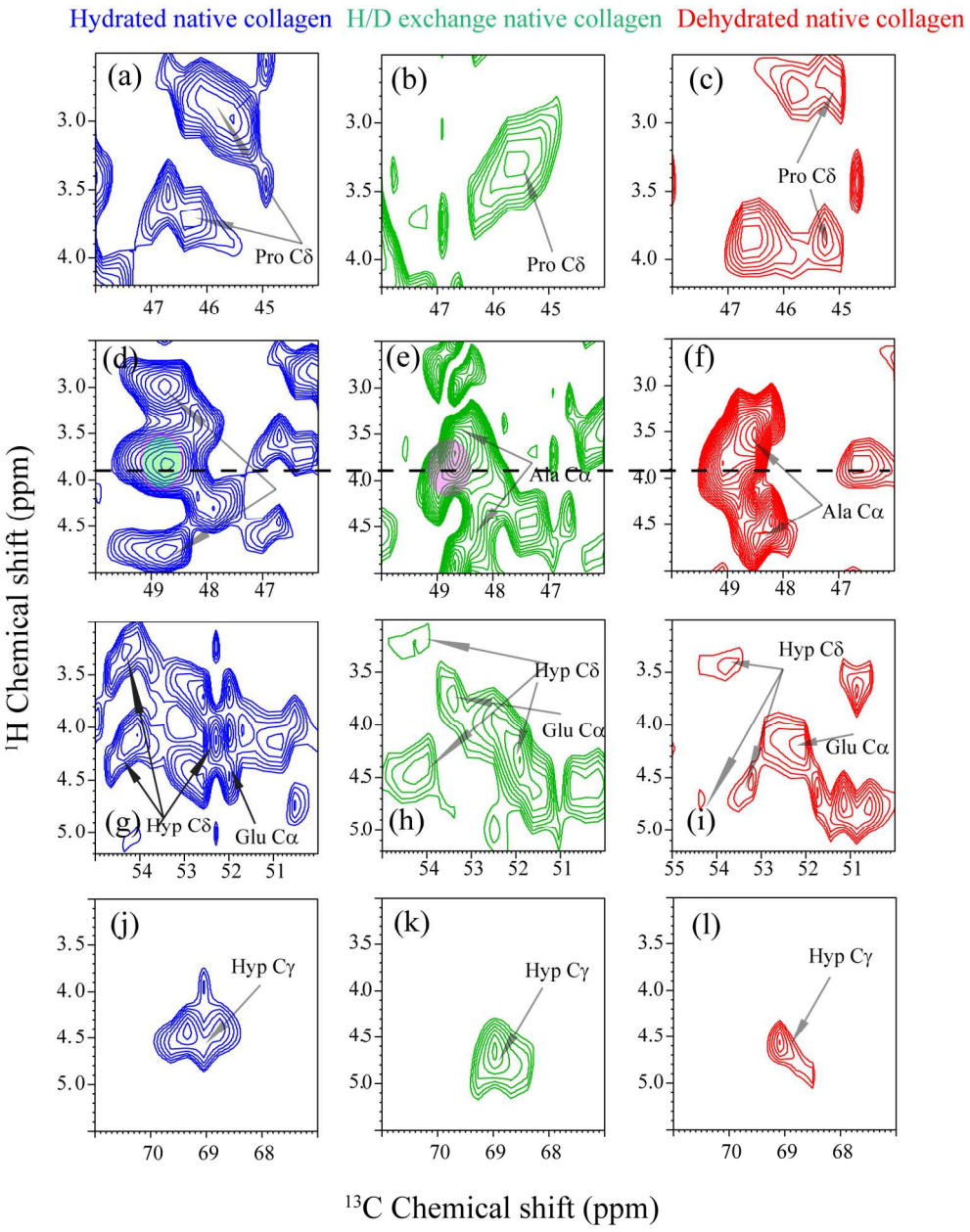

Figure S3. 2D $^{1}H/^{13}C$ HETCOR NMR spectra of hydrated native collagen, H/D exchanged and dehydrated native collagen showing changes in Hyp Cγ and Pro Cδ peak variations.

Simulation:

The evolution of 2D DQ – SQ peaks can be simulated for different dipolar coupling network. In order to infer corresponding changes at the structural level, we carried out SPINEVOLUTION simulation of DQ intensity as function of n_{rcpl}. The results of simulation along with dipolar coupling network around Hα are shown in **Figure S4**. For the simulations, we have used known X-ray structure of model collagen (PDB ID: 1CAG [10] and 1BKV:**Figure SI-5**). The detailed spin systems used for simulation are given in supporting information Table S10-17. The structure of triple helix consists of three chains hydrogen bonded to each other. Inter chain and intra chain simulation has been carried out to see the difference in dipolar coupling network. **Figure S4(a)** represent spin system network around $GlyCH_2$(a) showing protons within distance of 6Å. Distances of protons from different chains of triple helix are shown with different colors. The nearest protons from $GlyCH_2$(a) is $GlyCH_2$(b) at a distance of 1.86 Å from same residue/chain. It should be noted that Gly NH from same chain A is at a distance of 2.3 Å, and from another chain B Gly NH is at a distance of 2.5 Å (Figure 6(c)). This is followed by two more nearby protons, which are from the same chain corresponding to Pro CH and Pro CH_2δ at 2.69 Å and 2.7 Å respectively. Similarly other nearby proton network is shown in the **FigureS4(a) andS4(c)**. The dipolar coupling network corresponding to three chains (Figure 6(c)) is denser compared to single chain (Figure 6(a)). The corresponding DQ – SQ simulation for three chains of triple helix and single chain of triple helix are shown in **Figure S4(d) and S4(b)** respectively. The 2D DQ signal intensity for $GlyCH_2$(a) corresponding to three chains attains maximum value at the n_{rcpl} value of 2 whereas for single chain, it attain maximum at higher value of n_{rcpl} = 3. When dipolar-coupling network is dense, the maximum intensity of DQ – SQ evolution peak reaches at lower value of n_{rcpl}, as we observe in case of multiple chain simulation. Few other simulations were also carried out on model collagen peptide (PDB ID: 1BKV). This model collagen peptide shows sequence dependent conformational variation in collagen structure. In this model peptide we have chosen amino acids, which are outside of core in collagen triple helix. The simulation were carried out around Hα of Ala(13) CH, Arg(44)CH, and Leu(16)CH

corresponding to single and triple chains. The residues Ala, Arg and Leu reside outside the core of collagen triple helix. Simulation results show that there are subtle changes observed in single - and three – chain evolution. In the case of Ala CH, we found that other chain contribution shows its presence at distance of 4.0 Å, before that five nearest neighbor are present within distance of 3.0 Å. The evolution is mainly governed byintra chain protons network (Figure S4(e) and S4(g)). Hence, evolution of DQ – SQ curve is similar to inter and intra chain networks. In the case of Arg(44) CH maxima was observed around n_{rcpl}=4 and 3, in case of single chain and triple chain respectively (Figure **S4**(j) and **S4**(l)). This shift can be explained if we observe ^{1}H dipolar network as shown in Fi**gure S4(i, k)**. Three protons from other chain interfere within 3.5Å range, which changes initial build up of DQ – SQ evolution curve. In case of Leu(16), we observed that there are 8 nearby ^{1}H spins belonging to same chain as shown in **Figure S4(m, o).** Thus it has been observed that we get similar results in case of single chain and three-chain simulation **Figure S4 (n, p).** Thus we will not observed any difference in the evolution of Leu(13)CH peaks in the presence and absence of multiple chain. The above simulation indicates that DQ – SQ evolution of residues inside collagen triple helix core is more susceptible to changes in helix geometry.

Shifting of DQ-SQ peaks intensity values in native system such as collagen, was difficult to predict in case of multi-spin system. So instead of observing maximum value at corresponding n_{rcpl},we can curve-fit the DQ-SQ values as described in equation (1). The corresponding values of*b* will be directly proportional to the strength of dipolar coupling network in given system. We further probe above simulation of spin system by curve fitting by equation (2). This will gives different values of b and $1/t_{eff}$, which will tell how dipolar coupling network willvary in model spin system ((PDB ID: 1CAG [10] and 1BKV). **Figure S6(b,d),** clearly indicate that b values is quite sensitive to dipolar coupling network. We found that b values changes to 0.648 to 0.669 as effect of other chains were considered. Similar changes were also been observed in case of Ala(13) and Arg(46) **(Figure S6(f,h and g,l)).** This clearly establishes the fact that as b values changes, dipolar-coupling network becomes denser. It also

shows as dipolar coupling network changes **Figure S6 (a,e and i),** b values also changes as shown in **Figure S6 (b,f and j)**.

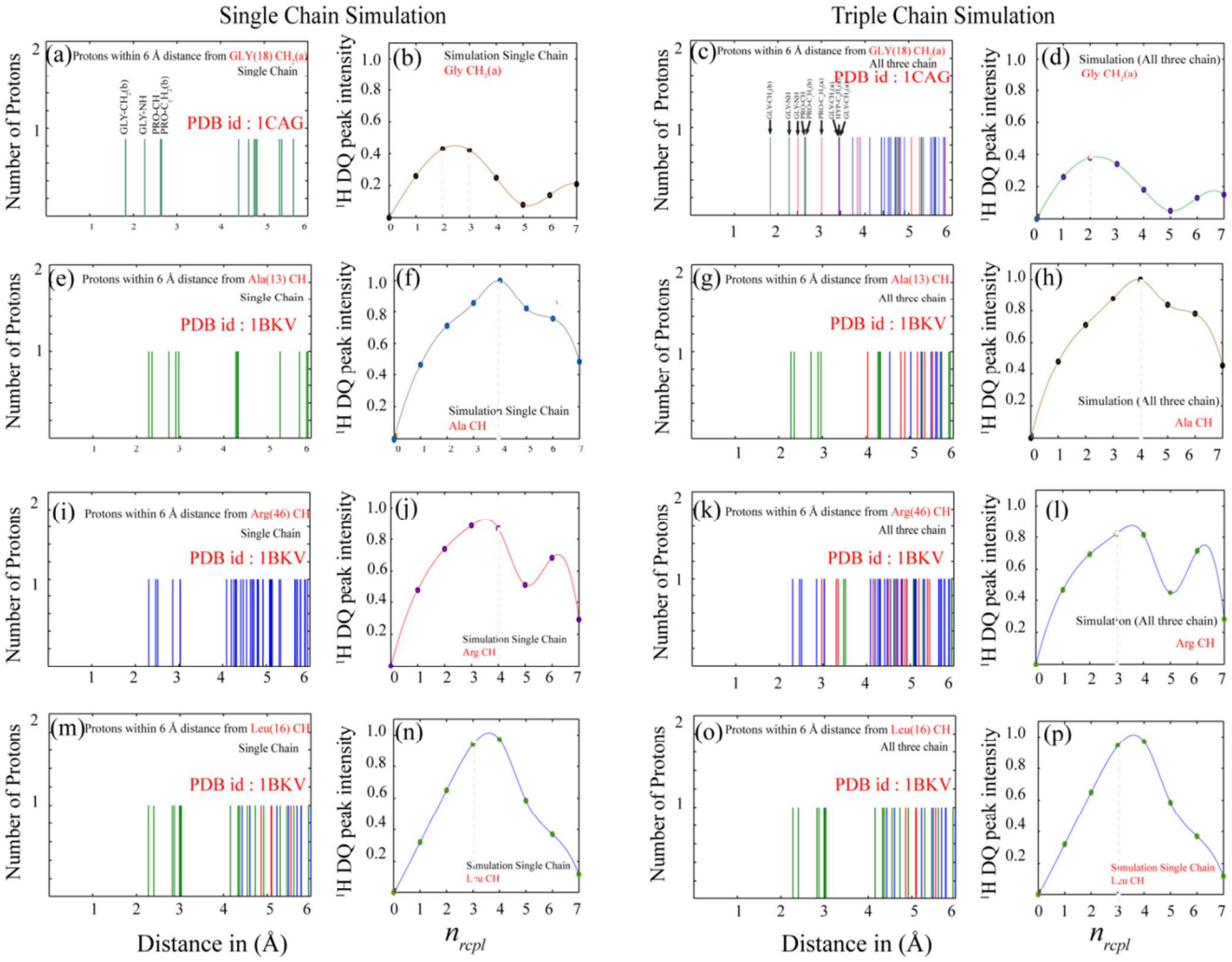

Figure S4: *^{1}H Dipolar coupling network (a and d) along with simulated DQ buildup intensity as a function of the number, n_{rcpl}, of POST-C7 elements used for the excitation and reconversion of double-quantum (DQ) coherence, corresponding to triple helix of collagen structure (PDB id: 1CAG) around Gly (18) CH_2 considering (a) and (b) of single chains, and (c) and (d) corresponds all three chains of collagen. (a) Protons corresponding to different chains within 6-Å distances are shown with different colors. Similarly for the another triple helix structure (PDB id: 1BKV), (e), (i) and (m) represents Protons corresponding to single chains within 6-Å corresponding to Ala(13) CH, Arg (43) CH, and Leu (16) CH. Corresponding to these spin system simulation of single chain was shown in (f), (j) and (n). For three chain triple helix proton spins within 6 Å distance and their simlations were represented in (g), (k), (o), (h), (l) and (p).*

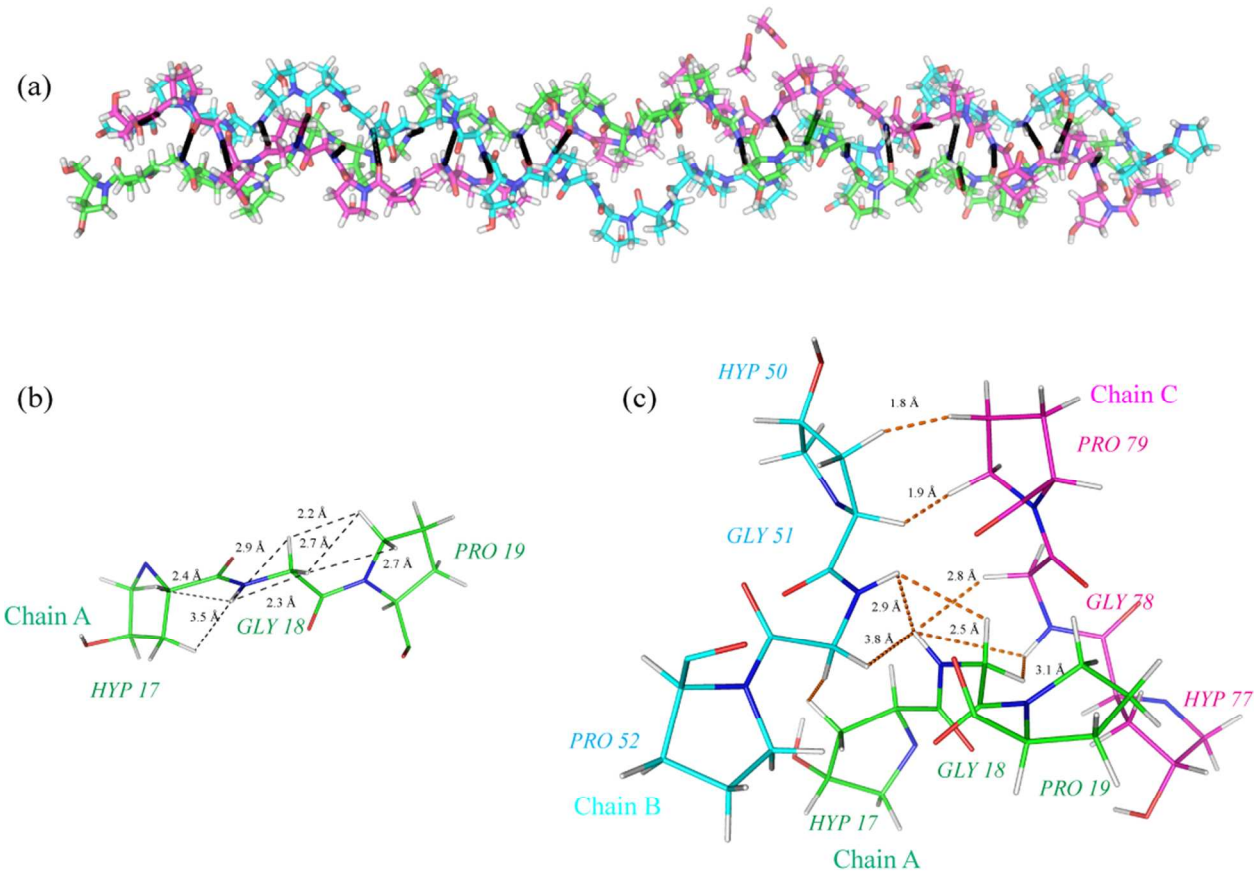

Figure S5. Representation of inter and intra-chain possible coupled network in collagen triple helix. (a) Pictorial representation of collagen triple helical structure. Ball and stick arrangement with NH $_{(GLY)}$ --- O=C $_{(Xaa)}$ hydrogen bondsform within triple helix. [PDB id: 1CAG]. (b) Representation of intra chain residue $^{1}H - ^{1}H$ pair within 3.5 Å from HN of Gly18 residue. (c) Representation of inter chain residue $^{1}H - ^{1}H$ pair within 3.5 Å from HN of Gly18 from Gly 51 and Gly 78 residues.

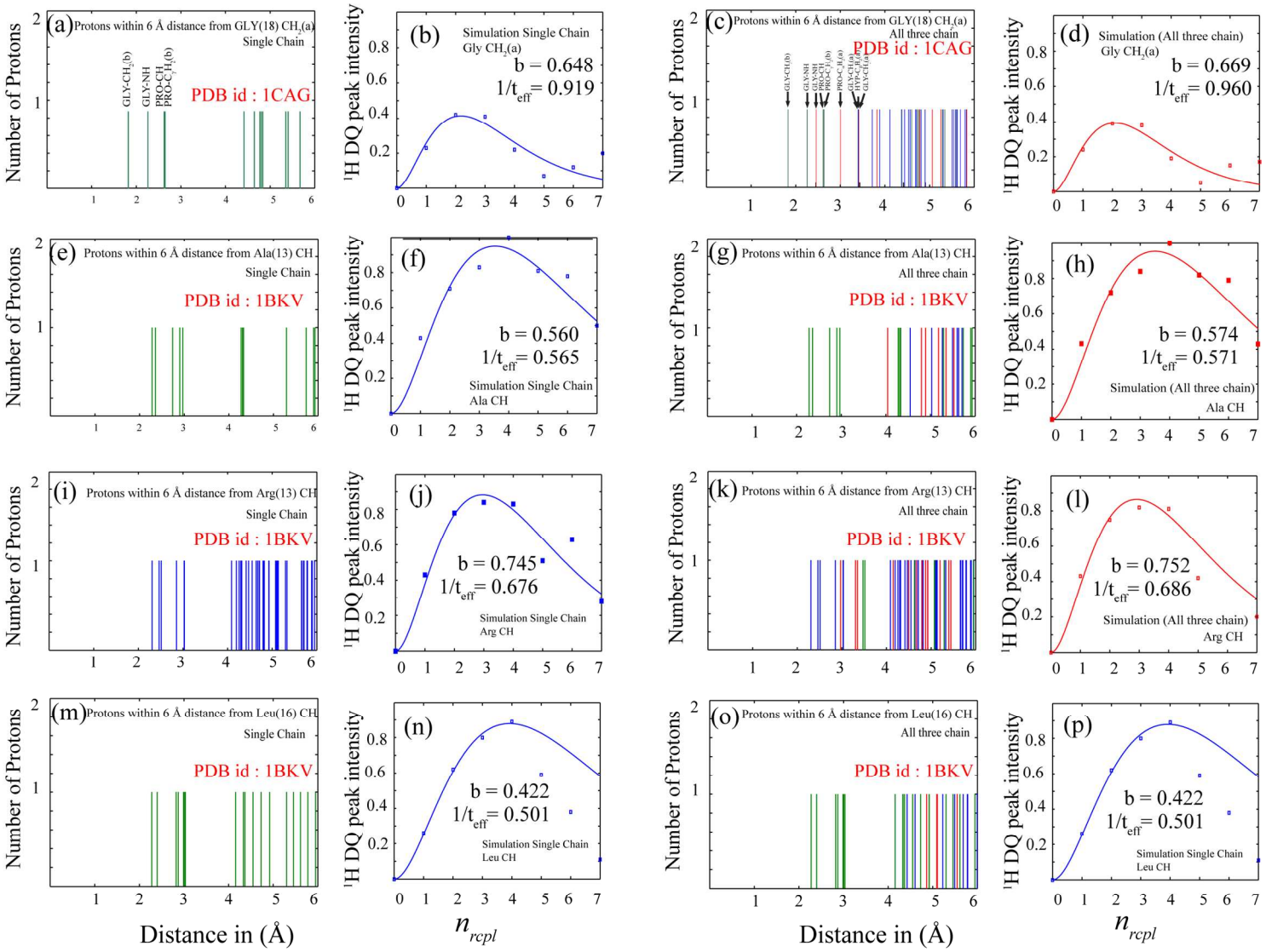

Figure S6: Simulation curves as mentioned in Figure S4 along with curve fit to the ^{1}H – ^{1}H DQ – SQ build up curve depicting dependence of parameter *b* with dipolar coupling network. It can be seen that parameter *b* is a strong indicator of local dipolar couplings around Hα.

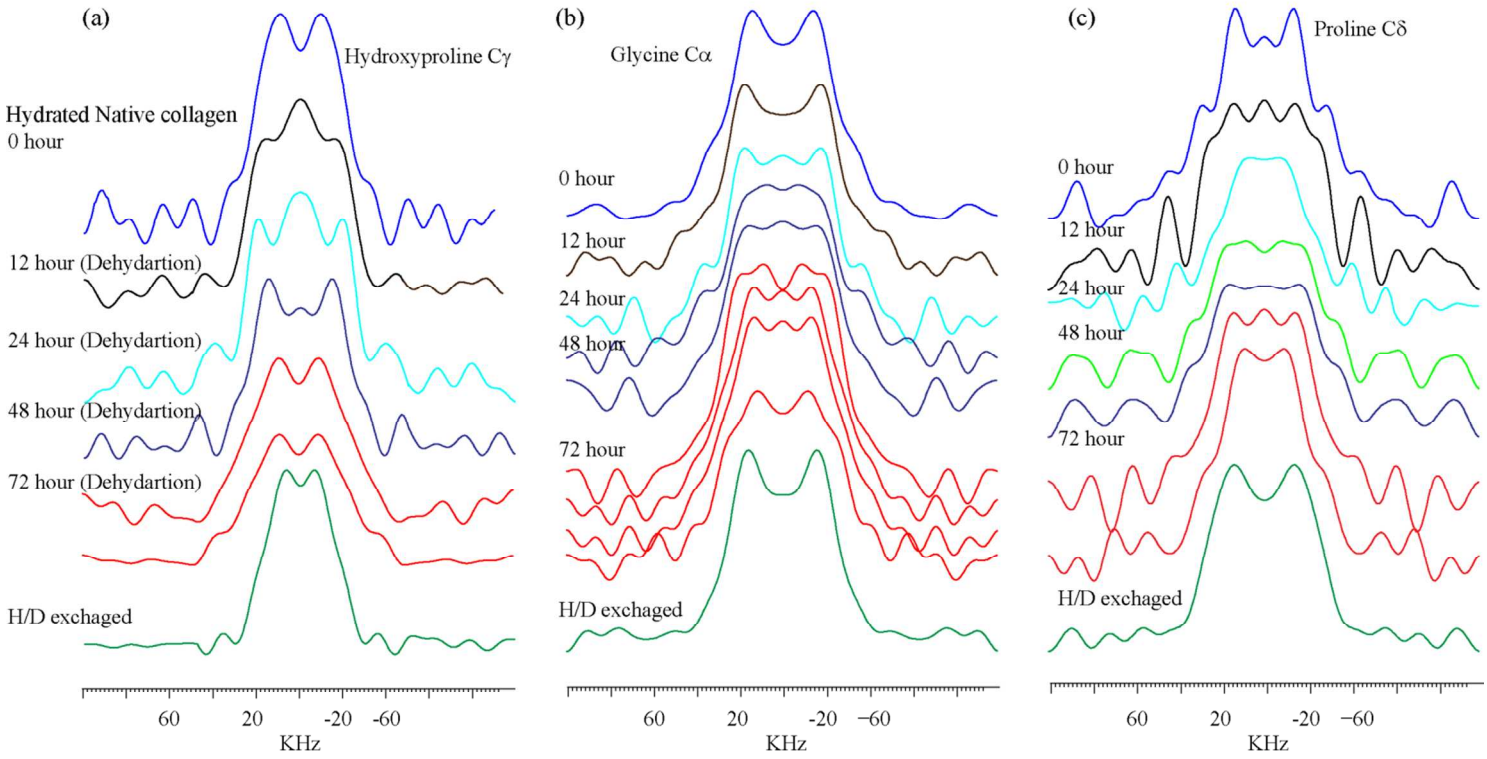

Figure S7.[1]*H WISE slice spectra of collagen samples with various degrees of dehydration and H/D exchanged for (a) Hydroxyproline*C_γ*(b) Glycine* C_α*(c) Proline*C_δ*. For recording WISE, we serially dehydrate the bone sample to check whether dehydration could change collagen side chain dynamics fast enough to affect DQ peak intensity? The WISE spectra qualitatively provide site-specific information about molecular dynamics of collagen protein.* [1]*H line width of collagen protein as function of dehydration clearly indicates that in case of hydroxyprolineCγ,glycine Cα and prolineXδ did not changes significantly. But in case H/D exchange native collagen sample we find that hydroxyprolineCγ*[1]*H line width changes up to 22% and 16 % compared to 72 hour dehydrated and hydrated native collagen sample. This change may occur due to change in hydrogen bonding network as hydroxyprolineCγ involves in hydrogen bonding network. For other peaks, we did not observe changes greater than 5 %.*

Table S2: Observed DQ peaks on ^{1}H DQ-SQ CRAMPS spectrum ofhydratednative collagen

S.No	Spin pair	*Sum of SQ freqs (ppm)*	*DQ freq (ppm)*
1	NH-CH	3.4 + 9.9	13.3
2	NH-CH	3.1+8.7	11.8
3	CH-NH	4.9+4.9	11.3
4	CH-NH	4.2+8.7	12.9
5	CH-NH	3.2+8.3	11.5
6	CH- NH	3.5+8.3	11.8
7	CH-NH	3.6+8.8	12.4
8	CH-NH	3.7+8.9	12.6
9	CH-NH	3.8+9.1	12.9
10	CH-NH	3.8+8.2	12.0
11	CH-NH	4.2+8.9	13.1
12	CH -NH	4.2+9.1	13.3
13	CH-NH	4.3+9.6	13.9
14	NH-CH	4.4+8.0	12.4
15	NH-CH	4.5+8.5	13.0
16	CH-NH	4.4+8.5	12.9
17	CH-NH	4.5+8.9	13.4
18	CH-NH	4.6+9.8	14.4
19	CH- NH	4.7+8.0	12.7
20	CH-NH	4.7+8.6	13.3
21	CH-NH	4.9+6.4	11.3
22	CH-NH	4.9+8.3	13.2
23	CH-NH	4.9+8.9	13.8
24	CH-NH	5.0+7.7	12.7
25	CH -NH	5.2+8.4	13.6

Table S3: Observed DQ peaks on ^{1}H DQ-SQ CRAMPS spectrum of H/D exchangednativecollagen

S.No	Spin pair	*Sum of SQ freqs (ppm)*	*DQ freq (ppm)*
1	NH-CH	2.8 + 9.8	12.6
2	NH-CH	2.7 + 9.6	12.3
3	CH-NH	2.9 + 8.1	11.0
4	CH-NH	2.9 + 7.1	10.0
5	CH-NH	3.1 + 9.1	12.2
6	CH- NH	3.2 + 8.2	11.4
7	CH-NH	3.5 + 8.9	12.4
8	CH-NH	3.7 + 8.2	11.9
9	CH-NH	3.8 + 9.3	13.1
10	CH-NH	3.9 + 7.4	11.3
11	CH-NH	3.9 + 8.8	12.7
12	CH -NH	4.0 + 8.0	12.0
13	CH-NH	4.1 + 7.4	11.5
14	NH-CH	4.2 + 8.6	12.8
15	NH-CH	4.4 + 8.6	13.0
16	CH-NH	4.5 + 7.8	12.3
17	CH-NH	4.7 + 7.1	11.8
18	CH-NH	4.7 + 6.3	11.0
19	CH- NH	5.1 + 6.6	11.7
20	CH-NH	5.0 + 7.0	12.0
21	CH-NH	5.1 + 8.1	13.2

Table S4: Observed DQ peaks on ^{1}H DQ-SQ CRAMPS spectrum of dehydrated native collagen

S.No	Spin pair	*Sum of SQ freqs (ppm)*	*DQ freq (ppm)*
1	NH-CH	2.9 + 8.1	11.0
2	NH-CH	2.9 + 8.8	11.7
3	CH-NH	2.9 + 7.1	10.0
4	CH-NH	3.1 + 9.0	12.1
5	CH-NH	3.0 + 8.5	11.5
6	CH- NH	3.2 + 9.1	12.3
7	CH-NH	3.5 + 8.5	12.0
8	CH-NH	3.4 + 9.2	12.6
9	CH-NH	3.2 + 7.9	11.1
10	CH-NH	3.5 + 7.5	11.0
11	CH-NH	3.4 + 8.6	12.0
12	CH -NH	3.8 + 7.4	11.2
13	CH-NH	3.6 + 8.7	12.3
14	NH-CH	3.8 + 8.3	12.1
15	NH-CH	4.0 + 9.2	13.2
16	CH-NH	4.0 + 8.5	12.5
17	CH-NH	4 .0 + 8.0	12.0
18	CH-NH	4.4 + 6.7	11.1
19	CH- NH	4.2 + 7.3	11.5
20	CH-NH	4.1 + 8.4	12.5
21	CH-NH	4.2 + 7.8	12.0
22	NH-CH	4.6 + 7.1	11.7
23	CH-NH	4.6 + 9.3	13.9
24	CH-NH	4.7 + 8.5	13.2
25	CH-NH	5.1 + 7.5	12.6
26	CH- NH	4.6 + 6.9	11.5
27	CH-NH	4.8 + 8.8	13.6
28	CH-NH	4.5 + 8.6	13.1
29	NH-CH	5.0 + 9.1	14.1

Table S5: Observed DQ peaks on ^{1}H DQ-SQ CRAMPS spectrum of type -1 collagen

S.No	Spin pair	*Sum of SQ freqs (ppm)*	*DQ freq (ppm)*
1	NH-CH	2.7 + 8.9	2.7
2	NH-CH	2.8 + 7.2	2.8
3	CH-NH	3.1 + 7.3	3.1
4	CH-NH	3.0 + 8.0	3.0
5	CH-NH	3.0 + 8.5	3.0
6	CH- NH	3.0 + 9.2	3.0
7	CH-NH	3.1 + 7.8	3.1
8	CH-NH	3.3 + 8.8	3.3
9	CH-NH	3. 4 + 7.8	3.4
10	CH-NH	3.5 + 9.3	3.5
11	CH-NH	3.6 + 9.2	3.6
12	CH -NH	3.5 + 8.7	3.5
13	CH-NH	3.8 + 7.5	3.8
14	NH-CH	3.8 + 8.6	3.8
15	NH-CH	3.8 + 8.1	3.8
16	CH-NH	4.1 + 8.0	4.1
17	CH-NH	4.1 + 8.6	4.1
18	CH-NH	4.3 + 8.0	4.3
19	CH- NH	4.4 + 9.1	4.4
20	CH-NH	4.7 + 7.7	4.7
21	CH-NH	4.5 + 7.3	4.5
22	NH-CH	4.6 + 8.5	4.6
23	CH-NH	4.9 + 7.4	4.9
24	CH-NH	4.7 + 9.5	4.7
25	CH-NH	5.0 + 8.8	5.0
26	CH- NH	4.8 + 8.7	4.8
27	CH-NH	5.3 + 7.0	5.3

Table S6: Curve fitting parameter of hydrated native collagen

SN	***b***	***$1/t_{eff}$***
1	*0.409*	*0.8743*
2	*0.4799*	*0.9387*
3	*0.4886*	*0.9216*
4	*0.3867*	*0.541*
5	*0.3018*	*0.6658*
6	*0.3669*	*0.6636*
7	*0.2856*	*0.5191*
8	*0.2281*	*0.4443*
9	*0.2809*	*0.5898*
10	*0.409*	*0.8743*
11	*1.018*	*0.9616*
12	*0.1486*	*0.2286*
13	*0.1204*	*0.1808*
14	*0.4136*	*0.6977*
15	*0.4536*	*0.7066*
16	*0.386*	*0.6344*
17	*0.4045*	*0.827*
18	*0.3609*	*0.6437*
19	*0.348*	*0.7282*
20	*0.617*	*0.9929*
21	*0.3549*	*0.8431*
22	*0.1592*	*0.2317*
23	*0.1911*	*0.2762*
24	*0.1534*	*0.2356*
25	*0.234*	*0.351*

Table S7: Curve fitting parameter of dehydrated native collagen

SN	***b***	***$1/t_{eff}$***
1	*0.5259*	*0.8026*
2	*0.6106*	*0.8634*
3	*0.7817*	*0.9589*
4	*0.7893*	*0.9384*
5	*0.2734*	*0.6457*
6	*0.3917*	*0.7612*
7	*0.5458*	*0.8682*
8	*0.479*	*0.7743*
9	*0.2953*	*0.5782*
10	*0.5154*	*0.8283*
11	*0.6505*	*0.8328*
12	*0.6044*	*0.7548*
13	*1.009*	*0.894*
14	*0.4302*	*0.6968*
15	*1.009*	*0.894*
16	*0.6905*	*0.7644*
17	*1.4*	*0.857*
18	*1.29*	*0.9923*
19	*1.497*	*1.023*
20	*0.6691*	*0.7442*
21	*0.6093*	*0.6786*
22	*0.8611*	*0.8099*
23	*0.7373*	*0.8502*
24	*0.6823*	*0.9033*
25	*0.5259*	*0.8026*
26	*0.4265*	*0..6684*
27	*.6612*	*0.8421*

Table S8: Curve fitting parameter of H/D exchanged native collagen

SN	***b***	$1/t_{eff}$
1	*0.8569*	*1.7073*
2	*0.6563*	*0.9817*
3	*0.6654*	*0.9601*
4	*0.34*	*0.7097*
5	*0.307*	*0.7014*
6	*0.3173*	*0.6339*
7	*0.679*	*0.9566*
8	*0.7922*	*0.8747*
9	*1.065*	*1.037*
10	*0.4466*	*0.737*
11	*0.3656*	*0.6136*
12	*0.5501*	*0.6842*
13	*0.791*	*0.9243*
14	*0.2492*	*0.4497*
15	*1.204*	*0.8663*
16	*0.4232*	*0.5982*
17	*0.2836*	*0.7556*
18	*0.5705*	*0.8747*
19	*0.28*	*0.5616*
20	*0.4937*	*0.777*
21	*0.3367*	*0.7465*

Table S9: Curve fitting parameter of type -1 collagen

SN	***b***	$1/t_{eff}$
1	*0.5114*	*0.8855*
2	*0.5859*	*0.8636*
3	*0.7498*	*0.9589*
4	*0.7514*	*0.9384*
5	*0.2622*	*0.6457*
6	*0.3755*	*0.761*
7	*0.5235*	*0.8622*
8	*0.4593*	*0.7742*
9	*0.2833*	*0.5782*
10	*0.4944*	*0.8283*
11	*0.6269*	*0.8328*
12	*0.5797*	*0.7548*
13	*0.9679*	*0.8939*
14	*0.4127*	*0.6968*
15	*0.9678*	*0.8939*
16	*0.6623*	*0.7644*
17	*1.343*	*0.8576*
18	*1.237*	*0.9922*
19	*1.436*	*1.023*
20	*0.6418*	*0.7422*
21	*0.5845*	*0.6787*
22	*0.826*	*0.81*
23	*0.7071*	*0.8512*
24	*0.6545*	*0.9033*
25	*0.5044*	*0.8026*
26	*0.576*	*0.7649*
27	*0.349*	*0.621*

Spin system coordinates used for simulations reported in figure S4 and S6.

Table S10. Glycine all chain simulation co-ordinate file

10.038	14.162	16.192	H2	(molecule A)
12.244	12.666	17.271	H1	(molecule A)
11.384	12.333	15.729	H3	(Observe)
13.195	10.807	15.1	H4	(molecule B)
10.318	12.986	13.168	H5	(molecule C)
14.07	11.1	16.657	H6	(molecule B)
4.994	13.891	16.419	H7	(molecule E)
8.58	12.269	17.668	H8	(molecule D)

Table S11. Glycine single chain simulation co-ordinate file

10.038	14.162	16.192	H2	(molecule A)
12.244	12.666	17.271	H1	(molecule A)
11.384	12.333	15.729	H3	(Observe)
13.195	10.807	15.19	H4	(molecule B)
14.07	11.12	16.657	H5	(molecule C)
4.994	13.891	16.419	H6	(molecule B)
9.247	15.655	17.572	H7	(molecule E)
15.117	10.375	13.931	H8	(molecule D)

Table S12. Arginine all chain simulation co-ordinate file

54.162	-4.703	12.16	H1	(molecule B)
55.971	-3.697	14.204	H2	(molecule A)
53.512	-3.871	14.212	H3	(observe)
54.872	-4.52	16.223	H4	(molecule A)
52.618	-5.226	16.559	H5	(molecule A)
53.184	-0.966	14.766	H6	(molecule C)
56.058	-5.386	13.589	H7	(molecule A)
54.91	-6.203	15.556	H8	(molecule A)

Table S13. Arginine single chain simulation co-ordinate file

54.162	-4.703	12.16	H1	(molecule B)
55.971	-3.697	14.204	H2	(molecule A)
53.512	-3.871	14.212	H3	(observe)
54.872	-4.52	16.223	H4	(molecule A)
52.618	-5.226	16.559	H5	(molecule A)
52.49	-6.863	11.28	H6	(molecule C)

56.058	-5.386	13.589	H7	(molecule A)
54.91	-6.203	15.556	H8	(molecule A)

Table S14. Alanine all chain simulation co-ordinate file

44.557	-0.684	12.327	H1	(molecule B)
44.38	-3.233	9.906	H2	(molecule A)
44.342	-3.571	12.162	H3	(observe)
44.136	-1.527	10.35	H4	(molecule a)
43.978	-0.391	16.922	H5	(molecule A)
45.773	-2.128	9.997	H6	(molecule B)
73.77	-19.442	-8.26	H7	(molecule C)
43.261	0.031	14.194	H8	(molecule D)

Table S15. Alanine single chain simulation co-ordinate file

44.557	-0.684	12.327	H1	(molecule B)
44.38	-3.233	9.906	H2	(molecule A)
44.342	-3.571	12.162	H3	(observe)
44.136	-1.527	10.35	H4	(molecule a)
43.978	-0.391	16.922	H5	(molecule A)
45.773	-2.128	9.997	H6	(molecule B)
47.191	-6.786	11.853	H7	(molecule C)
43.261	0.031	14.194	H8	(molecule D)

Table S16 .Leusine all chain simulation co-ordinate file

52.192	-2.265	6.051	H1	(molecule A)
52.996	-6.201	5.845	H2	(molecule B)
51.698	-4.528	6.675	H3	(observe)
51.744	-4.365	3.855	H4	(molecule A)
52.324	-2.387	8.483	H5	(molecule A)
53.885	-4.24	4.659	H6	(molecule A)
51.355	-2.692	4.322	H7	(molecule A)
53.961	-2.511	6.592	H8	(molecule A)

Table S17. Leusine single chain simulation co-ordinate file

52.192	-2.265	6.051	H1	(molecule A)
52.996	-6.201	5.845	H2	(molecule B)
51.698	-4.528	6.675	H3	(observe)
51.744	-4.365	3.855	H4	(molecule A)
52.324	-2.387	8.483	H5	(molecule A)
53.885	-4.24	4.659	H6	(molecule A)
51.355	-2.692	4.322	H7	(molecule A)
53.961	-2.511	6.592	H8	(molecule A)

References:

(1) Vinogradov, E.; Madhu, P. K.; Vega, S., High-Resolution Proton Solid-State NMR Spectroscopy by Phase-Modulated Lee–Goldburg Experiment. *Chem. Phys. Lett.* **1999,** *314*, 443-450.

(2) van Rossum, B. J.; Förster, H.; de Groot, H. J. M., High-Field and High-Speed CP-MAS13C NMR Heteronuclear Dipolar-Correlation Spectroscopy of Solids with Frequency-Switched Lee–Goldburg Homonuclear Decoupling. *Journal of Magnetic Resonance* **1997,** *124*, 516-519.

(3) Ladizhansky, V.; Vega, S., A Method for Measuring Heteronuclear (1H−13C) Distances in High Speed MAS NMR. *J. Am. Chem. Soc.* **2000,** *122*, 3465-3472.

(4) Hohwy, M.; Jakobsen, H. J.; Eden, M.; Levitt, M. H.; Nielsen, N. C., Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence. *J. Chem. Phys* **1998,** *108*, 2686-2694.

(5) Elena, B.; de Paëpe, G.; Emsley, L., Direct Spectral Optimisation of Proton–Proton Homonuclear Dipolar Decoupling in Solid-State NMR. *Chem. Phys. Lett.* **2004,** *398*, 532-538.

(6) Sakellariou, D.; Lesage, A.; Hodgkinson, P.; Emsley, L., Homonuclear Dipolar Decoupling in Solid-State NMR Using Continuous Phase Modulation. *Chem. Phys. Lett.* **2000,** *319*, 253-260.

(7) Brown, S. P.; Lesage, A.; Elena, B.; Emsley, L., Probing Proton−Proton Proximities in the Solid State: High-Resolution Two-Dimensional 1H−1H Double-Quantum CRAMPS NMR Spectroscopy. *Journal of the American Chemical Society* **2004,** *126*, 13230-13231.

(8) Lesage, A.; Sakellariou, D.; Hediger, S.; Eléna, B.; Charmont, P.; Steuernagel, S.; Emsley, L., Experimental Aspects of Proton NMR Spectroscopy in Solids Using Phase-Modulated Homonuclear Dipolar Decoupling. *J. Magn. Reson.* **2003,** *163*, 105-113.

(9) Schmidt-Rohr, K.; Clauss, J.; Spiess, H. W., Correlation of Structure, Mobility, and Morphological Information in Heterogeneous Polymer Materials by Two-Dimensional Wideline-Separation NMR Spectroscopy. *Macromolecules* **1992,** *25*, 3273-3277.

(10) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M., Crystal and Molecular Structure of a Collagen-Like Peptide at 1.9 A Resolution. *Science* **1994,** *266*, 75-81.