**Supplementary Information**

Figure S1: Onset stress $\sigma_0$ of nonlinearity in the shear moduli of vimentin networks (left) and linear modulus $G_0$ as a function of loading rate. Dashed lines indicate average values, shaded areas indicate an interval of one standard deviation. Unlike the peak modulus and rupture strain (figure 2) both $G_0$ and $\sigma_0$ are independent of loading rate.
Figure S2: Typical strain vs. stress curve of a vimentin network ($\dot{\sigma} = 0.1 s^{-1}$) obtained using the stress ramp protocol. Network rupture is defined as having taken place when the strain increases by at least 5 orders of magnitude between successive measurements. Filled symbols show data before rupture, empty symbols show data after rupture.

Figure S3: Loading rate dependent contribution to differential modulus $K'$ as a function of normalized stress. The contribution is calculated by subtracting loading rate dependent nonlinear rheology data (figure 2) from loading rate independent data (figure 1). At low applied stress, all data superimposes, and the contribution is negligible. At applied stresses exceeding $\sigma^*$, loading rate dependent deviations from stress pulse data are observed, with slower loading rates corresponding to larger deviations.
Figure S4: Simulation of the response of affinely deforming polymer networks under applied shear. (a) Initially isotropic filament ensembles are generated by placing filaments at random positions and orientations and applying a homogeneous shear strain. The force per filament is calculated using an analytical expression for the force-extension relation of an inextensible semiflexible polymer. (b) The evolution of the resulting storage modulus (black line) is in approximate agreement with the differential storage modulus measured through the stress-pulse protocol of 1mg/ml vimentin (red squares). Discrepancies between simulation and experiment are seen at $\sigma \approx \sigma_0$, which is likely due to slight network nonaffinity as agreement is observed at higher vimentin concentrations of 2mg/ml (blue triangles, taken from [3]). (c) The force per filament remains below 1pN for all filaments in the linear regime ($\sigma < \sigma_0$) and at the onset of stiffening ($\sigma \approx \sigma_0$). At the yield stress ($\sigma \approx \sigma$) the force per filament exceeds 1pN for a small proportion ($\approx 15\%$) of filaments.
Figure S5: Frequency sweeps of vimentin networks in different crosslink regimes, averaged (N=3), prepared at concentrations of 1mg/ml, strain amplitude 0.5%. Transiently crosslinked networks have comparable values of $G'$ and $G''$ to permanently crosslinked networks, indicating that the glutaraldehyde crosslinking does not significantly alter the network architecture or dynamics. When crosslinking is suppressed, both $G'$ and $G''$ are notably lower across all frequencies, which is likely a consequence of their increased mass per length, resulting in a coarser network at equivalent concentrations. Further evidence for this is presented below in figure S6.

Figure S6: Linear (a) and nonlinear (b) rheological response of tailless vimentin networks at different concentrations and total filament length per volume, $\rho$. Where the concentration of tailless vimentin is identical to wild type vimentin the linear response is significantly weaker and the degree of stiffening is significantly lower. When both networks have identical $\rho$ their linear response is identical, but the degree of stiffening remains low, clearly demonstrating that crosslinking is the main determinant of reduced stiffening.
References