Decomposing dynamical couplings in mutated scFv antibody fragments into stabilizing and destabilizing effects

Azhagiya Singam Ettayapuram Ramaprasad[1], Shahid Uddin[2], Jose Casas-Finet[3] and Donald J. Jacobs[1]*

[1]Department of Physics and Optical Science, University of North Carolina at Charlotte; [2]Formulation Sciences, MedImmune Ltd., Cambridge, UK; [3]Analytical Biochemistry Department, MedImmune LLC, Gaithersburg, Maryland.

ABSTRACT: Conformational fluctuations within scFv antibodies are characterized by a novel perturbation-response decomposition of molecular dynamics trajectories. Both perturbation and response profiles are stratified into stabilizing and destabilizing conditions. The linker between the VH and VL domains exhibits the dominant dynamical response by being coupled to nearly the entire protein, responding to both stabilizing and destabilizing perturbations. Perturbations within complementarity-determining regions (CDR) induce rich behavior in dynamic response. Among many effects, stabilizing any CDR loop in the VH domain triggers a destabilizing response in all CDR loops in the VL domain and vice versa. Destabilizing residues within the VL domain are likely to stabilize all CDR loops in the VH domain, and, when these residues are not buried the CDR loops in the VL domain are also likely to be stabilized. These effects, described by shifts in normal mode characteristics, initiate a propensity for dynamic allostery with possible functional implications in bispecific antibodies.

Introduction

A single chain fragment variable (scFv) is the smallest antibody molecule that binds to an antigen. These scFv antibodies are important to medicinal applications such as development of immunotoxins, therapeutic gene delivery, and development of bifunctional proteins for therapeutic purposes.[1-4] The structure of the scFv antibody is comprised of two variable regions of heavy (VH) and light (VL) chains connected by a flexible linker. The antigen binding site in a scFv antibody is formed by six loops (three from VH and three from VL domains) that define the complementarity-determining regions (CDRs). Much attention and effort has been placed on conducting trial and error structure/function studies to engineer antibodies to have specific binding targets and stability characteristics.[5-8] Unfortunately, scFv antibodies often are associated with stability problems that limit their utility.

Stability can be improved through site-directed mutagenesis, but often at the expense of modifying the specificity and selectivity of antibody antigen interactions. A naïve rational design strategy is to mutate one or more residues that are remote from binding sites. With improved stability, the structural alignment of the mutated and native proteins is likely to be close, suggesting function will be preserved using the structure/function paradigm. However, specificity and selectivity depend on protein dynamics. Stabilizing mutations can alter correlated motions between residues without conformational change.[9] More generally, the possible mechanisms responsible for the sensitivity in structure function relationships arguably parallel those governing allosteric effects.[10-11].

Allostery[12] refers to the process where binding of an effector at an allosteric site changes the binding affinity of a ligand at a distal binding site.[11,13,14-16] The ensemble nature of allosteric response encompasses a continuum of cooperative processes that involve protein dynamics with, and without, persistent conformational change.[17] *Dynamic allostery*[18-23] occurs when dynamic fluctuations are altered without conformational change. A plausible mechanism is through a change in the characteristics of normal modes of vibration about an average structure.[18] Although normal modes typically adapt to a perturbation on protein structure, the presence of allostery requires a functional change. Therefore, the dynamic fluctuations at a binding site (its mobility) must change in sufficient magnitude to modify specificity or selectivity. This means the intrinsic dynamical property of a protein[24-25] informs on the potentiality for measuring an allosteric response, but which outcome is observed depends on the location of the perturbation and binding details of the protein-effector and protein-ligand.

There is also evidence that allosteric response is sensitive to mutations in terms of dynamical properties based on computational results[26] and in terms of allosteric regulation based on large-scale experimental structure/function studies.[27] Here dynamical coupling between all residue pairs is decomposed to establish a baseline for the propensity of dynamic allostery, and to obtain a higher order signature of protein dynamics to help elucidate stability and functional characteristics across protein mutants.

How dynamical fluctuations change upon external perturbation to protein structure has been extensively investigated using a variety of computational methods.[28-31] A common rationale for these methods is to apply a local rigidifying perturbation to mimic ligand binding at a site, and subsequently identifying a response at any other site having a significant change in dynamics (or flexibility). The perturbations are applied as a scan along the backbone of a protein. Perturbation-response methods fall mainly into three schemes that employ an elastic network model (ENM),[32-35] a distance constraint model (DCM)[26] or molecular dynamics (MD) simulation.[36-37] Other methods investigate dynamic correlations from a covariance matrix involving atomic positions based on a MD trajectory for an unperturbed protein,[29-30] and, subsequently, apply principal component analysis (PCA).[38-39] Most commonly ENMs are employed because long-range effects are captured well by low-frequency vibrational modes.[25, 40]

A recent trend is to include more atomic details about dynamic fluctuations and interactions.[40] Examples include an ensemble based atomistic ENM where a consensus network is derived from multiple input structures[38]. Another approach makes use of the equivalence of eigenvectors from the Hessian matrix[41-42] representing normal modes of vibration to PCA-modes from the covariance matrix. Importantly, the covariance matrix is proportional to the inverse Hessian matrix.[38, 41-43]. Covariance matrices from MD simulations are also employed to transform (as a matrix multiplication) a local external force perturbation on the protein into dynamic correlations as a linear response.[30-31] Another recent perturbation-response method runs $n+1$ independent MD simulations on a protein when it is unperturbed and when it is subjected to n different applied local perturbations in a scan over n residues.[36-37] Unfortunately, this method is computationally expensive, requiring an independent MD simulation for each perturbation variant.

Six bispecific scFv antibodies are studied as a continuation to previous investigations[9, 26] where 5 mutants differ by 1 to 4 point mutations from the native scFv. Notably, each mutant increases stability and retains similar antigen binding affinity. Recently it has been observed bispecific scFv antibodies exhibit cooperative communication between binding loops across domains,[44-45] and the functional role of linkers[46-47] is known to be important, making this system particularly interesting to study.

Herein, a novel perturbation-response method is implemented based on an effective Hessian matrix obtained from the inverse of the covariance matrix that is obtained from all-atom MD simulation of an unperturbed protein. This effective Hessian matrix represents an ensemble based elastic network with long-range couplings that capture collective quasi-harmonic motions. Local perturbations are applied to the elastic network as a scan, yielding a new Hessian matrix per perturbation. Both stabilizing and destabilizing perturbations are considered. Normal modes of vibration are calculated by exact diagonalization for each new Hessian matrix. Dynamical coupling quantifies changes in the mean squared fluctuation (MSF) of C_α atoms that are distal from the perturbed site. Dynamic couplings for stabilizing and destabilizing perturbations are decomposed into stabilizing and destabilizing response.

Methods

The set of 100 *ns* all-atom MD trajectories from prior studies were reused to sample 2000 conformation for a scFv antibody and each of five stabilizing mutations at room temperature; 6 MD trajectories in total.[9, 26] Additionally, one representative conformation was selected from each MD trajectory within the most and least populated clusters based on RMSD clustering analysis. Starting from these two structures the same protocols were followed to run 100 ns MD simulations using the GROMACS 5.1.2 package[48] resulting in three independent MD trajectories for native scFv and for each of the 5 mutants.

A $3N \times 3N$ covariance matrix, Q, is constructed for each MD trajectory by tracking the aligned x, y, z coordinates of the C_α atoms in the protein.[49] Each of the six scFv fragments have $N = 238$ residues. A Hessian matrix can be formally obtained from Q^{-1}. However, the smallest eigenvalues of Q dominate its inverse, which results in random noise in the Hessian matrix. This noise is removed by filtering the spectral decomposition of Q. Eigenvectors and eigenvalues of Q are first calculated where $|n\rangle$ is the n-th eigenvector with eigenvalue λ_n for an ordered indexing such that $\lambda_{n+1} > \lambda_n$. The eigenvectors are PCA modes and the eigenvalues define variances in collective atomic positions along a PCA mode direction. The PCA modes are normalized and form a complete set such that $\langle m|n\rangle = \delta_{mn}$ and $1 = \sum_n |n\rangle\langle n|$. If an inverse is not possible due to a zero eigenvalue, a minimum value of λ_{min} can be defined such that $\lambda_n \to \max(\lambda_n, \lambda_{min})$. However, this flooring operation was not needed because the smallest eigenvalue of Q posed no difficulty, such that the inverse of Q exists in all cases, albeit noise is present due to small $\lambda_n$values.

Noise decorrelation: From spectral decomposition it follows that $Q = \sum_n |n\rangle\, \lambda_n\, \langle n|$ and $Q^{-1} = \sum_n |n\rangle\, \lambda_n^{-1}\, \langle n|$. However, numerical error is present in the estimate for λ_n and the direction of its corresponding mode $|n\rangle$. Assuming this error is $\Delta\lambda$ for all modes, there will be a critical index such that $\lambda_c \sim \Delta\lambda$ for which the PCA modes with $n \geq c$ will cause Q^{-1} to be physically meaningless. In contrast, these modes represent small amplitude motions in Q that beget negligible contribution. To restore physical significance of Q^{-1} the random correlations are removed by making the substitution $\lambda_n \to \langle\lambda\rangle_{n\geq c} = \frac{1}{3N-c}\sum_{n=c}^{3N}\lambda_n$ over the $n \geq c$ subspace. This transformation maps fine scale motion into a degenerate subspace, while keeping the trace of the covariance matrix invariant. The cutoff index, c, is determined by the percent of variance to be decorrelated. Coverage of 10% works well, meaning the total variance of all modes with $n \geq c$ is 10% of the trace of the covariance matrix. Coverage between 5 to 20% yields similar results. All data shown here is based on decorrelating noise at 10% coverage, meaning 90% of the protein dynamics is retained.

Effective Hessian matrix: The decorrelated Q^{-1} is used to define a Hessian matrix, $H_o = RT\, Q^{-1}$, representing an effective ENM involving C_α atoms, similar to an anisotropic network model.[43] Note that RT is thermal energy. Importantly, H_o is derived from quasi-harmonic PCA modes that reflect all-atom dynamics of the protein simulated by MD. As such, H_o has long-range non-zero matrix elements not found in ENMs that focus on nearest neighbor interactions within 15 Angstroms. Examples of the covariance matrix and effective Hessian matrices at different noise decorrelation levels are shown in supporting information **Figure S1**. It is worth noting that different H_o having cutoffs between 0% to 20% coverage maps back to virtually the exact same covariance matrix. Subsequently, H_o is perturbed by H_p to obtain $H = H_o + H_p$.

To calculate MSF for all C_α atoms, PCA modes and corresponding variances for the perturbed system are obtained by diagonalizing H, noting that $Q_p = RTH_p^{-1}$ formally. The eigenvalues of H_p are given as ω_n, where $\omega_n < \omega_{n+1}$ and $\lambda_n = RT/\omega_n$, and ω_n are proportional to the frequency of vibration of normal modes. When working with a Hessian matrix, it is common practice to define a pseudo inverse by removing the lowest six eigenvalues that are formally zeros, representing 3 rigid body translations and 3 rigid body rotations. However, for effective Hessian matrices truncating the lowest six eigenvalues does not allow Q to be reproduced from H^{-1}. Physically, the lowest eigenvalues of H_p correspond to the largest variance in collective motion.[41] As such, the flooring process of $\omega_n \to \max(\omega_n, \omega_{min})$ is applied with $\omega_{min} = 0.9/\lambda_{max}$ and λ_{max} is the largest eigenvalue of the covariance matrix of the unperturbed system, enabling Q_p to be defined. This establishes a physically reasonable threshold that restricts the variance in any PCA mode of Q_p not to exceed 11.1% of the maximum PCA mode variance in the unperturbed system.

Perturbation Characteristics: Six types of local perturbations are considered. Each perturbation adds a set of springs between certain pairs of C_α atoms relative to a reference C_α atom at residue p. All added springs are assigned the same spring constant, k_p. The unperturbed protein's average coordinates defined in the covariance matrix are used to set the natural length of each spring. Adding these springs do not change the equilibrium point of the structure, but conformational fluctuations about the equilibrium point are modified. Among six types of perturbations, three footprints denoted as triad, star and ball are considered. A triad footprint adds three springs to the elastic network to connect pairs of C_α atoms located at $\{(p, p+1), (p-1, p), (p-1, p+1)\}$. At the N- and C-termini only one of the three springs is added. A star footprint adds springs to the elastic network to connect C_α atom at residue p to all C_α atoms located within a radius of R_p from the C_α atom at p. A ball footprint adds springs to the elastic network to connect all pairs of C_α atoms that fall within R_p relative to the C_α atom at p. Schematic pictures of each footprint type are shown in **Figure S2** of supporting information.

The size of the star and ball footprints can be adjusted by R_p. In this study, a range between 8Å to 12Å is considered. Stabilizing and destabilizing perturbations can be generated by positive and negative k_p respectively. To make comparisons of different size star and ball perturbations less dependent on size, k_p is scaled as $k_p = k/N_p$ where N_p is the number of independent vibrational modes when the subsystem perturbed is viewed as a separate unit.

Perturbation Response Matrix: For each perturbation, a spectral decomposition of H_p yields new eigenvectors and eigenvalues. The eigenvalues are transformed into PCA variances, $\{\lambda_{pn}\}$ where p is the residue number for the reference location of the perturbation and n is a PCA mode index where $\lambda_{p,n+1} > \lambda_{p,n}$. The MSF at each residue is calculated as $MSF(p,r) = \sum_n \lambda_{p,n} V_n^2(r)$ where r is the residue where the mean square fluctuation is observed, and $V_n(r)$ is the r-th component of the n-th PCA mode. Compared to the unperturbed protein $MSF(p,r) - MSF(0,r)$ defines a perturbation-response matrix (PRM), and $p = 0$ defines the unperturbed case. Since distal changes are of interest, local effects are removed by excluding matrix elements whenever the distance between C_α atoms at p and r are 15Å or less. For these pairs of residues the PRM elements are set to zero. For stabilizing perturbations ($k > 0$) the PRM is denoted as $M_s(p,r)$. Likewise, $M_d(p,r)$ denotes the PRM for destabilizing perturbations ($k < 0$). Letting the variable Y denote perturbation type s or d, $M_Y(p,r)$ denotes a PRM for either stabilizing or destabilizing perturbations where the rows define a perturbation scan, and the columns define the location of the response.

A decrease or increase in MSF at a given residue in response to a perturbation is respectively referred to as stabilizing or destabilizing. The PRM is divided into stabilizing and destabilizing parts as $M_Y = M_{Ys} + M_{Yd}$ such that all negative matrix elements in M_Y are the same as in M_{Ys}, while all other matrix elements in M_{Ys} are zero. Positive responses are similarly tracked by M_{Yd}. A four-state classification scheme in the form of YX is defined as (ss, sd, dd, ds) where Y and X refer to the perturbation and response respectively. In summary, the PRM for stabilizing perturbations is decomposed into $M_{ss}(p,r)$ and $M_{sd}(p,r)$ and for destabilizing perturbations the PRM is decomposed into $M_{ds}(p,r)$ and $M_{dd}(p,r)$.

Response and Perturbation Profiles: A response profile for an entire scan of stabilizing perturbations across a protein is given by $R_{sX}(r) = \sum_p M_{sX}(p,r)/C_s$, with a normalization constant defined as $C_s = \sum_{pr}[\,|M_{ss}(p,r)| + M_{sd}(p,r)\,]$. The $R_{ss}(r)$ profile is for stabilizing response, and the sum $p_{ss} = \sum_r R_{ss}(r)$ gives the percent of residues that undergo a stabilizing response for stabilizing perturbations. Similarly, $R_{sd}(r)$ is the response profile for residues that are destabilized over a scan of stabilizing perturbations, and p_{sd} gives the percent of residues that have a destabilizing response. Similar equations are defined for destabilizing perturbations, yielding the $R_{ds}(r)$ and $R_{dd}(r)$ response profiles, and associated p_{ds} and p_{dd} that respectively give the percent of residues having a stabilizing (rigidify) or destabilizing (increase flexibility) response due to a destabilizing perturbation. Perturbation profiles are also defined. For example, $P_{ss}(r) = \sum_r M_{ss}(p,r)/C_s$ quantifies where stabilizing perturbations will likely cause a stabilizing response. The profiles $P_{sd}(r)$, $P_{ds}(r)$ and $P_{dd}(r)$ are also similarly defined. Note that total percentages (p_{ss}, p_{sd}, p_{ds} p_{dd}) obtained from perturbation and response profiles are equal because summing over r first and then p second, or p first and then r second, gives the same results.

Cross-sections of information within the PRM are conveyed by $M_{XY}(p_o,r)$, which tracks the response at each residue, r, for a perturbation applied at p_o. If the interest is to determine which residues should be perturbed that will lead to a response at r_o, this information is conveyed as $M_{XY}(p,r_o)$. The profiles defined above show an average distribution across the entire protein. Other profiles can be constructed for a targeted region, such as for stabilizing response for stabilizing perturbations in the CDR2H loop that is given by $R_{ss}(r|CDR2H) = \sum_{p\,\in H2} M_{ss}(p,r)/C_s$. Summing this response profile over all residues in the protein gives the total response that comes from only perturbed residues in the H2 loop. Dividing this sum by p_{ss} results in a relative percent for how much the stabilizing perturbations in the H2 loop causes a response throughout the protein. Considering different targets enables dynamical couplings to be efficiently analyzed.

Results and Discussion

Miller *et al* experimentally identified a set of mutant scFv fragments with similar activity.[6] The structure has a heavy VH and light VL domain connected by a flexible linker at residues 116-132. Five mutants with increased stability are considered here, ordered with increasing melting temperature: V56G, P104D, 2M (S16E & S177L), 3M (S16E, V56G & S177L) and 4M (S16E, V56G, P104D, S177L). Locations of all mutant sites are shown on the structure in supporting information **Figure S3**, including the labeling of key structural features. Dynamic coupling is investigated through changes in MSF caused by external perturbations with response/perturbation profiles compared.

The change in mobility at each residue is calculated over a scan of perturbations applied at each residue, in turn, one at a time. Both stabilizing and destabilizing perturbations are considered. For stabilizing perturbations, elastic forces with a characteristic spring constant, k, create a clamping effect to suppress conformational fluctuations at the perturbation site. Changing the sign of the characteristic spring constant from positive to negative yields a destabilizing perturbation. In this case, conformational fluctuations that squeeze together C_α atoms within a perturbation site are suppressed; otherwise the C_α atoms are wedged apart with greater force as their separation increases farther from their equilibrium separations. The intrinsic elasticity of a protein must overpower destabilizing forces to prevent a local (or global) unfolding event.

At distal residues from a perturbation site, a positive or negative change in MSF relative to the unperturbed protein respectively yields a destabilizing or stabilizing response. This process results in a perturbation-response matrix (PRM) for stabilizing and destabilizing perturbations, shown in **Figure 1** for the native protein. The PRM patterns from stabilizing and destabilizing perturbations are not anti-symmetric. For all proteins, stabilizing perturbations within the complementarity-determining regions (CDR-loops) lead to overall increases in mobility within the protein, indicated by red horizontal strips.

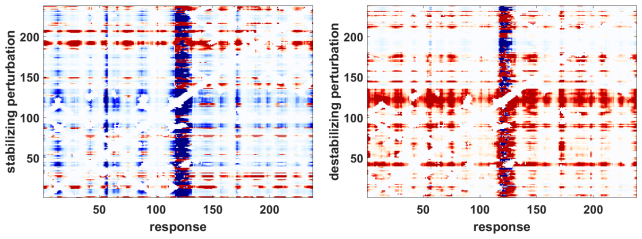


Figure 1: Native PRM for stabilizing (left) and destabilizing perturbations (right). Red and blue colors respectively indicate an increase and decrease in MSF, and white indicates no change.

The PRMs for stabilizing and destabilizing perturbation for all six proteins are compared in supporting information in **Figures S4** and **S5** respectively. Interestingly, similar PRM patterns group into three separate pairings such that V56G is similar to native, P104D is similar to 4M with a stronger response to perturbation, and 2M is similar to 3M with a weaker response. These groupings are similar to groupings by binding affinities.[6] However, the rank order of these PRM pairs by strength is not the same as the rank order by binding affinities or melting temperature.[6]

Global Mobility Response: Summing MSF over all residues in a protein is equal to the trace of the covariance matrix that defines total mobility. Similarly, summing MSF changes due to a perturbation over all residues yields the change in total mobility. The response profiles for total mobility are shown in **Figure 2** for

stabilizing and destabilizing perturbation on all six proteins. These profiles show where perturbations must be applied on a protein to increase or decrease total mobility. The profiles for perturbations that are stabilizing (**Figure 2a**) or destabilizing (**Figure 2b**) respectively pair up in the same way the PRMs pair up. Namely, in order from weakest to strongest response, the pairs (2M, 3M), (V56G, Native) and (P104D, 4M) exhibit similar profile patterns.

Comparing the profile patterns of a protein from stabilizing and destabilizing perturbations reveals an approximate anti-symmetric character. Across the six proteins, linear correlation coefficients between profiles for stabilizing and destabilizing perturbations fall in the range -0.65 < R < -0.79. Consequently, whenever a stabilizing perturbation causes total mobility to increase, a destabilizing perturbation is likely to cause mobility to decrease and vice versa. In supporting information, **Table S1** lists correlation coefficients between all protein pairs for respective quantities that decompose changes in total mobility into stabilizing or destabilizing response to either a stabilizing or destabilizing perturbations. Linear correlation coefficients between respective quantities across all pairs of proteins fall in the range 0.70 < R < 0.94 for stabilizing perturbation and 0.55 < R < 0.94 for destabilizing perturbation.

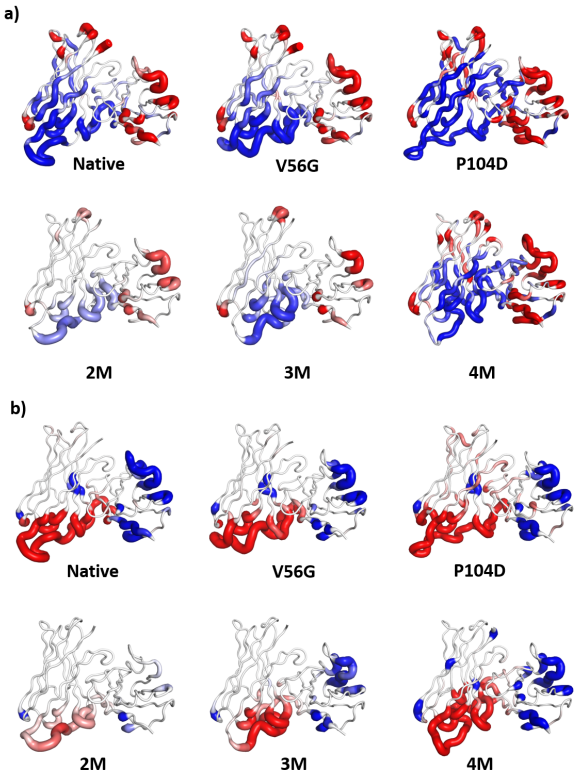


Figure 2: Change in total mobility is mapped onto structure due to a) stabilizing and b) destabilizing perturbations. Blue and red colors show reduced and increased total mobility respectively. White indicates virtually no change. These results are for a ball footprint with $R_p = 10$Å and $k_o = 10\ kcal/(mol\ \text{Å}^2)$.

Notwithstanding variations in magnitude, profile patterns for total mobility are well conserved for a given type of perturbation, which is reflecting the protein fold as expected. Since the perturbations considered in this work alter conformational fluctuations at a perturbation site without changing equilibrium positions, it may be possible that a perturbation that disrupts the fold could create a response pattern that deviates far from that observed here. Assuming the protein remains folded and functional as it interacts with other molecules the qualitative anti-symmetric character of global response should be quite general, although changes in total mobility do depend on properties of the perturbations. In particular the largest differences appear between ball and triad footprints. Triad results are shown in supplemental information in **Figure S6**.

Mean Response Distribution: Dynamic response is decomposed into average response profiles for the cases ss, sd, dd and sd. General trends on how the response depends on the size and strength of a ball footprint is shown in supplemental information in **Figure S7, S8 and S9**. The ball radius is explored in the range from 8Å to 12Å, as the characteristic spring constant governing strength of a perturbation is explored in a range from 1 to 20 $kcal/(mol\ \text{Å}^2)$. A robust pattern emerges from two independent trends. First, as the perturbation radius (R_p for ball or star) increases, the average response profile is enhanced for ss and dd cases, but diminished for sd and ds cases. This indicates that as R_p increases the average response becomes more compliant with a perturbation. This effect is likely a general result, as it is intuitive that a larger footprint will more effectively drive a protein to respond in the direction of the perturbation. This trend is consistent with a previous finding that highly localized rigidifying perturbations induce flexibility at distal locations in similar spirit to the Le Chatelier's principle.[9]

For residues outside the flexible linker between the VH and VL domains, an increase in perturbation strength increases destabilizing response and decreases stabilizing response. Combining size and strength dependence shows that ss, sd, dd and ds response profiles are enhanced when (size, strength) is respectively (large, small), (small, large), (large, large) and (small, small). Results for the ball footprint with R_p= 10Å and $k_o = 10\ kcal/(mol\ \text{Å}^2)$ provide a representative example from which all main conclusions can be inferred. The relative magnitudes in the response for dynamic coupling for the six proteins are summarized in **Tables 1** and **2** for stabilizing and destabilizing perturbations respectively.

The quantity TR_o listed in **Tables 1, 2** is the trace of the covariance matrix (total mobility) of the unperturbed protein. The listed quantity $\langle|\Delta TR|\rangle$ is the average change in trace upon perturbation based on a uniform scan over the entire protein, which characterizes the overall magnitude of the dynamic response. It is apparent that the rank ordering of the observed pairings (2M, 3M), (V56G, Native) and (P104D, 4M) follows the rank ordering in $\langle|\Delta TR|\rangle$. The magnitude of dynamic response linearly correlates with TR_o well (R= 0.96 for both stabilizing and destabilizing perturbations), indicating that the magnitude of dynamic response is essentially proportional to the global mobility of a protein.

Table 1: Size and relative distribution characteristics of dynamic response to stabilizing perturbations for all six proteins ordered from top to bottom with the least to greatest stability. The units for trace quantities are in Å^2.

Protein	TR_o	$\langle\|\Delta TR\|\rangle$	rel%S	%SP	%ss	%sd
Native	506	73.6	100.0	39.8	67.9	32.1
V56G	513	57.9	78.6	36.5	65.8	34.2
P104D	685	134.4	182.7	36.8	61.3	38.7
2M	251	18.0	24.4	37.8	62.8	37.2
3M	295	27.2	37.0	41.1	69.4	30.6
4M	528	95.2	129.4	39.5	64.1	35.9

Table 2: Size and relative distribution characteristics of dynamic response to destabilizing perturbations for all six proteins ordered from top to bottom with the least to greatest stability. The units for trace quantities are in Å^2.

Protein	TR_o	$\langle\|\Delta TR\|\rangle$	rel%D	%DP	%ds	%dd
Native	506	111.5	100.0	60.2	14.7	85.3
V56G	513	100.7	90.3	63.5	7.3	92.7
P104D	685	230.6	206.8	63.2	8.8	91.2
2M	251	29.6	26.5	62.2	7.9	92.1
3M	295	39.0	34.9	58.9	14.0	86.0
4M	528	146.0	131.0	60.5	12.1	87.9

The $\langle|\Delta TR|\rangle$ data is normalized relative to the native protein so that each protein is assigned a relative percent rel%S or rel%D for stabilizing or destabilizing perturbations respectively. Likewise, %SP and %DP give percentages for the relative size of dynamic

response due to stabilizing or destabilizing perturbations. The %ss and %sd give relative percentages for how much a stabilizing perturbation leads to a stabilizing or destabilizing response. Similarly, %ds and %dd provide relative percentages for a destabilizing perturbation. The information in **Tables 1, 2** taken together indicate that the overall characteristics of dynamic response in all six proteins are markedly similar, with the notable difference that the native, 3M and 4M proteins have nearly twice the stabilizing response to destabilizing perturbations than the V56G, P104D and 2M proteins. No "rule of thumb" was found to correlate affinity or stability to a dynamical coupling signature of some type, such as total number of perturbation sites that trigger long-range dynamic coupling, or those at the interface, or within functional loops.

Linker Communication: The linker has the greatest response to perturbation compared to anywhere else in the protein, indicating a dynamical communication pathway exists between the linker and the majority of the protein. Mobility in the linker can increase or decrease due to stabilizing or destabilizing perturbations (e.g. **Figure S8**). Interestingly, change in mobility within the linker as perturbation strength increases responds in the opposite way to all (or nearly all) other residues in the protein. For the ss and ds cases, the effectiveness of a perturbation to stabilize the linker decreases as perturbation strength increases. Conversely, the effectiveness to destabilize the linker decreases. Furthermore, as perturbation strength increases the magnitude of the response in the linker decreases regardless of perturbation or response type. Glycine-serine rich linkers are known to be important in the scFv fragments which helps in maintaining their activity.[50-51] Flexibility in this linker is also associated with top ranked PCA modes that allow its high mobility to be captured. These results elucidate the important role that the linker plays in scFv fragments.

Variable Response Profiles: Linear correlation coefficients for mean response profiles between mutants to the native protein are given in **Table 3**. The tabulated data indicates there is overall similarity in response profiles between native and any mutant (especially for the case sd) but variation in profile shape (not relative magnitudes) is significant (15 out of 40 cases $R < 2/3$) with mutant 2M deviating from native the most. The response profiles for each protein are also linearly correlated to MSF, with R-values listed in **Table 4**. Only 2 out of 48 cases has $R < 2/3$, and for 43 cases $R > 3/4$, indicating that dynamic response tracks MSF well. Correlation between all protein pairs is given in **Table S2**, where the response in 2M deviates the most from all other systems.

Table 3: Correlation coefficients for linear correlation in response profiles between mutant and native with and without the linker.

	ss		**sd**		**dd**		**Ds**	
	Linker	No Linker	Linker	No Linker	Linker	No Linker	Linker	No Linker
V56G	0.81	0.63	0.91	0.93	0.90	0.75	0.73	0.48
P104D	0.98	0.45	0.90	0.93	0.93	0.88	0.97	0.63
2M	0.53	0.36	0.68	0.87	0.84	0.39	0.32	0.48
3M	0.93	0.21	0.67	0.88	0.82	0.52	0.95	0.16
4M	0.84	0.53	0.79	0.82	0.84	0.63	0.75	0.52

Table 4: Correlation coefficients for linear correlation in response profiles between protein and its MSF with and without the linker.

	ss		**sd**		**dd**		**ds**	
	Linker	No Linker	Linker	No Linker	Linker	No Linker	Linker	No Linker
Native	-0.98	-0.89	0.93	0.68	0.86	0.90	-0.96	-0.80
V56G	-0.97	-0.88	0.75	0.65	0.84	0.90	-0.95	-0.82
P104D	-0.98	-0.94	0.84	0.73	0.81	0.85	-0.97	-0.81
2M	-0.98	-0.88	0.89	0.76	0.89	0.70	-0.90	-0.81
3M	-0.97	-0.82	0.77	0.71	0.84	0.93	-0.94	-0.68
4M	-0.99	-0.91	0.75	0.56	0.85	0.80	-0.97	-0.83

Previous studies show that regions with high mobility in proteins play a vital role in allosteric regulation.[24, 52] Note that destabilizing response for stabilizing perturbation (sd) is strongly correlated with or without the linker. The sd case is often attributed to function by the rationale that flexibility is induced at a binding site in response to an allosteric effector. Taken together, this result suggests all mutations have similar sd response as that of the native protein to conserve function. Interestingly, the correlation in response profiles between native and all mutants is high for the ss and ds cases, but excluding the linker leads to a substantial drop in correlation. This result indicates that the linker plays a dominant role in stabilizing the protein.

Mean response profiles for each protein are shown in **Figure 3**, where different scales for color rendering are used for each case (i.e. ss, sd, dd, ds) to obtain maximum contrast, but the same scale is used across all proteins to allow variations between proteins to be discerned. The mean response profiles within a classification show similarity and variation across the proteins. Similar trends in **Figure 3** show the ss and dd cases as having a strong response in the linker and in the CDR loops of the VH domain. In addition, strong destabilizing response to stabilizing perturbations (case sd) occurs in the linker and all CDR loops, suggesting that an increase in mobility within CDR loops help facilitate protein function. For the ds case, the linker always destabilizes in response to a destabilizing perturbation.

Antisymmetric patterns are not present between ss and dd cases or sd and ds cases. While the sd case is the most conserved, the ds case exhibits most variation across proteins than any other category. Mean ds response profiles for 2M and 4M mutants respectively show least and greatest response, and exhibit the most deviation on opposite ends of the spectrum compared to all other proteins. Upon close inspection, numerous small variations are observed in each protein across corresponding categories, even between pairs of proteins sharing a similar PRM.

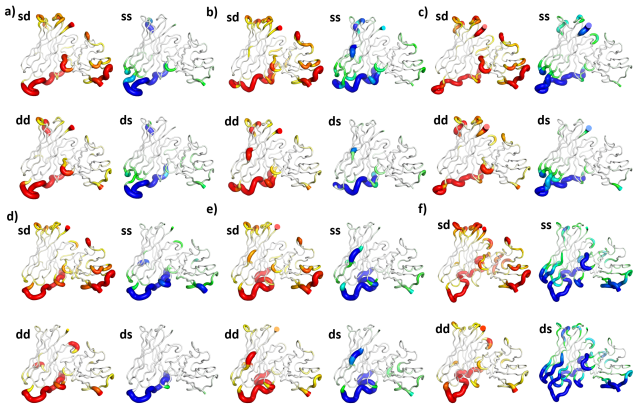


Figure 3: Mean response profiles mapped onto structure for a) native, b) V56G, c) P104D, d) 2M, e) 3M and f) 4M systems. The color scheme of red, orange, yellow orders destabilizing response from largest to smallest. The color scheme of blue, cyan, green orders stabilizing response from largest to smallest. White represents virtually no response.

CDR Loop Communications: Each CDR loop is in turn targeted with stabilizing and destabilizing perturbations. Their dynamic response profiles are shown in **Figure 4** for the native protein. Response profiles for all six CDR loop perturbation targets for all proteins are given in supporting information **Figures S10 – S15**. These profiles are expected to be functionally relevant since antigens interact with CDR loops.

As seen in **Figure 4**, dynamic response spans the VH and VL domains. For example, if CDR1H in VH is perturbed, a signal transfers to the VL domain via the linker and the interface between domains. The conditional response profiles for the categories ss, sd dd and ds differ depending on which CDR loop is perturbed and how. Interestingly, in all the systems a destabilizing perturbation in any CDR loop in the VL domain triggers a response in all CDR loops of the VH domain. Perturbation in the

CDR1 or CDR3 loops in the VH domain or any CDR loop in the VL domain induces stabilizing response in the CDR2H loop in all the systems. Furthermore, stabilizing perturbation to the CDR1H loop triggers a destabilizing response to all CDR loops in the VL domain. It is worth noting that a stabilizing perturbation to the CDR2H or CDR3H loops also trigger a destabilizing response to all CDR loops in the VL domain only in P104D and 4M systems. Stabilizing perturbations of most CDR loops trigger a dynamic response in the CDR2H loop, which is the longest loop present in the scFv fragments. These results suggest that increase in mobility facilitates recognition of antigen binding. In previous studies it was shown that binding of antigen to CDR-3 loop is responsible for the rearrangement of VH and VL domains.[53]

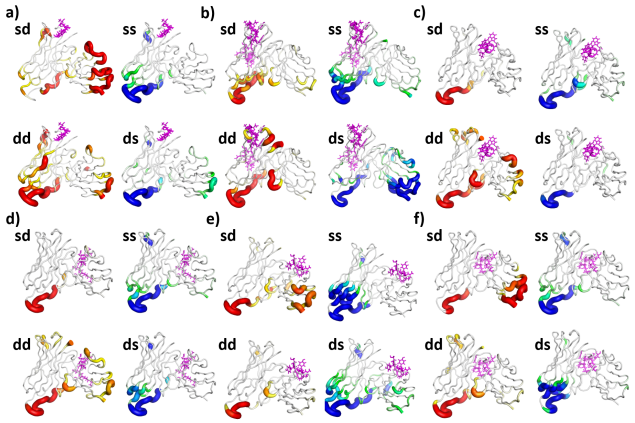


Figure 4: Mean response mapped onto structure for perturbation targets at the CDR loops (a) CDR1H, b) CDR2H, c) CDR3H, d) CDR1L, e) CDR2L, f) CDR3L. Magenta color highlights targeted CDR loops; otherwise the color scheme is the same as in Figure 3.

Perturbation Profiles: Perturbing a single residue can stabilize some residues and destabilize others. A mean perturbation profile differentiates which residues trigger a large or small response for the ss, sd, dd and ds categories. Common features among the scFv fragments are shown in **Figure 5** by averaging the mean perturbation profiles over all six proteins. Interestingly, hydrophobic residues in the VH/VL domains are respectively more/less compliant. Increased compliance means stabilizing perturbations yield more distal stabilization and likewise destabilizing perturbations yield more distal destabilization (e.g. ss and dd cases). The exposed hydrophobic residues are more compliant than buried hydrophobic residues. In addition, stabilizing perturbations on buried hydrophobic residues in the VH domain trigger less destabilization.

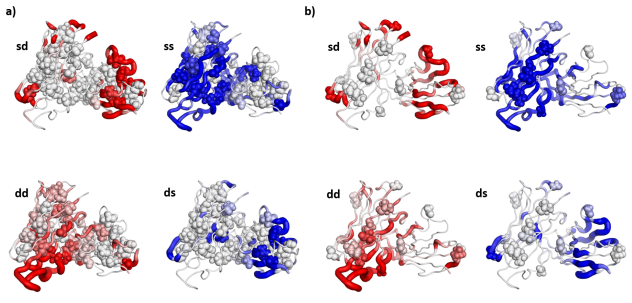


Figure 5: A representative mean perturbation profile is mapped onto native structure for ss, sd, dd, ds cases. Balls show locations of hydrophobic residues that are a) buried and b) exposed. Blue and red colors respectively show residues that trigger a stabilizing or destabilizing response. White color indicates no response.

Upon inspection, PRM patterns for all six scFv fragments (e.g. **Figures S4** and **S5**) exhibit overwhelming likelihood for residues along a row to either have stabilizing or destabilizing response. As such, perturbation profiles for the ss and sd cases are nearly orthogonal, meaning whichever case has the largest non-zero value, the other case is zero or negligible. Similarly, dd and ds cases are nearly orthogonal. A notable exception in these scFv fragments is that the linker often has a contra-response compared to all other residues (as noted above), which prevents perfect orthogonality. Nevertheless, ss and sd (or dd and ds) perturbation profiles can be overlaid on protein structure to show in one rendering which residues have greater stabilizing or destabilizing allosteric response.

In the same format for total mobility presented in **Figure 2**, the long-range (not local) mean perturbation profiles are shown in **Figure 6** for all systems. It can be seen that a cluster of residues in the beta sheet of the VH domain decreases mobility upon a stabilizing perturbation. The size of these clusters and spatial distribution in strength varies across the proteins, whereas the location of residues that trigger a destabilizing response is well conserved. These results suggest that the conserved pattern in destabilizing response is important to support function in these proteins.

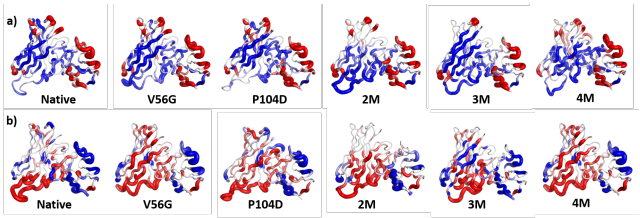


Figure 6: Ribbon diagrams for different scFv fragments are colored by a) stabilizing b) destabilizing perturbation profiles. Blue and red colors respectively show residues that trigger net stabilizing or destabilizing response. White color indicates no response.

As shown in **Table 5** only moderate to weak correlations are observed between mutant proteins and native with or without the linker, indicating that perturbation profiles are modulated by mutations. Moreover, for ss, sd and ds cases there is weak correlation to MSF (typically $|R| < 0.2$). Furthermore, correlations between mean perturbation profile of each mutant to MSF (while weak) is markedly larger, ranging from ($0.47 < R < 0.66$). It is clear that perturbation profiles do not track MSF, although response profiles do. These results indicate each mutant responds in a similar way to sites that are sensitive to perturbations, but the effectiveness of sites to affect a response differ across the mutants.

Table 5: Correlation coefficients for linear correlation in mean perturbation profiles between mutant and native with and without the linker.

	ss		**sd**		**dd**		**ds**	
	Linker	No Linker	Linker	No Linker	Linker	No Linker	Linker	No Linker
V56G	0.72	0.7	0.79	0.83	0.78	0.76	0.69	0.70
P104D	0.75	0.75	0.71	0.75	0.85	0.80	0.73	0.73
2M	0.38	0.50	0.84	0.86	0.70	0.67	0.26	0.34
3M	0.60	0.62	0.72	0.75	0.68	0.65	0.32	0.33
4M	0.58	0.59	0.70	0.73	0.76	0.67	0.53	0.55

Functionally Important Perturbation Sites: Residues that trigger distal response in CDR loops are expected to play an important role in stabilizing scFv fragments without losing activity. Therefore, a targeted response for each CDR-loop is separately considered to arrive at six conditional perturbation profiles shown in **Figure 7** for the native protein. These results highlight the most effective places to perturb the native protein to obtain a response in the CDR loop of interest. The same conditional stabilizing and destabilizing perturbation profiles for all mutant proteins are respectively shown in supporting information **Figures S16** and **S17**.

Results shown in **Figures 7**, **S16** and **S17** indicate stabilizing any CDR loop in the VH domain triggers a destabilizing response in any CDR loop in the VL domain and vice versa. In addition, stabilizing CDR2L or CDR1H triggers a destabilizing response in

any other CDR loop. Destabilizing any CDR loop in the VH domain triggers a destabilizing response in any other CDR loop. In general, many residues within CDR loops can be perturbed to trigger a stabilizing or destabilizing response in other CDR loops.

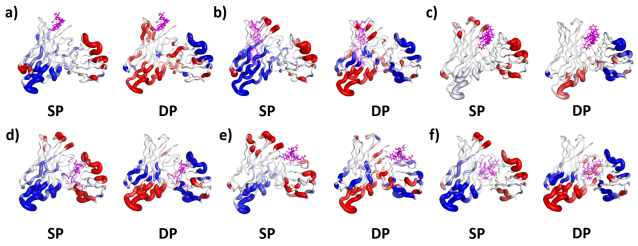


Figure 7: The ribbon diagrams of the native scFv fragment are colored by their average CDR3L that trigger response in different CDR loops a) CDR1H b) CDR2H c) CDR3H, d) CDR1L, e) CDR2L and f). Blue and red colors respectively show residues that trigger a stabilizing or destabilizing response for stabilizing perturbation (SP) or Destabilizing perturbation (DP). White color indicates no response and magenta color highlights the CDR loop.

Distinctive Perturbation Sites: To determine if certain sites have any distinctive characteristic if they are perturbed, the protein is decomposed into structurally and functionally important regions. Regions are classified in different ways. One classification assigns a residue to either the VH or VL structural domain or the interface between the two domains or the linker. Outside of the linker, residues are also classified as exposed, buried or transitional. Specifically, mean solvent accessible surface area (mSASA) with standard deviation was calculated for each residue using GROMACS over all MD simulations per protein. The distribution for mSASA is bimodal. A residue is buried when its mSASA plus its standard deviation is less than the dividing line between modes. It is exposed if its mSASA minus its standard deviation is greater than the dividing line, and transitional otherwise. Combining these two classifications defines 10 distinct regions in a protein for which every residue is in one and only one grouping, such as VH-buried, VH-exposed, VH-transitional, etc. In addition, the six CDR loops are of functional interest, yielding a total of 16 regions of interest.

A targeted perturbation is applied to each region of interest, and the ratio of mean response derived from a targeted perturbation to mean response from the entire protein (c.f. **Figure 3**) is calculated for regions of interest to arrive at a coarse-grained PRM. A ratio close to 1 indicates no distinctive characteristic. A ratio less than ½ or more than 2 is recorded as a signal, as shown for each protein and for the ss, sd, dd and ds cases in supporting information **Figures S18**, **S19** and **S20**. Consensus signals across all scFv fragments are shown in **Figure 8** when the p-value of a block is less than 0.0062 (i.e. requirement of 2.5 standard deviations away from the null hypothesis that there is not a signal).

From **Figure 8** salient trends can be inferred. Stabilizing perturbations in the linker have propensity to stabilize the entire protein. Perturbing the linker is unlikely to destabilize VH residues or residues in the interface. Destabilizing a buried residue in the VH domain likely yields a stabilizing dynamical response within VH, which extends to the entire protein if the residue is exposed. Destabilizing any interface residue is likely to stabilize residues in the VH domain, which extends to the entire protein if the residue is not buried. Destabilizing any residue in the VL domain is likely to stabilize all CDR loops in the VH domain, and if the residues are not buried it is likely that CDR loops in the VL domain will also be stabilized. Stabilizing CDR2L has propensity to destabilize CDR1H and CDR2H, while destabilizing loop CDR2L has propensity to stabilize CDR1H and CDR2H. Stabilizing CDR1H has a propensity to destabilize all loops in the VH domain, especially loop CDR2L. It is worth noting that the presence of positive and negative cooperativity between CDR loops have been experimentally observed in monoclonal antibodies.[54]

Perturbing at the mutation sites (16, 56, 104, 177) leads to differing results across all mutants, however stabilizing perturbations at residue-16 yields a consensus (data not shown) for the propensity to stabilize the entire protein. This means that most mutation site locations that both stabilize the protein and maintain function do not standout with any special characteristics. Importantly, the perturbation profile is hyper sensitive to mutation for the ss, sd, dd, ds classifications, leading to a mix of positive and negative cooperativity effects on the CDR loops within the same domain and across domains. These perturbation sites are numerous, and thus likely to facilitate non-specific binding of solute molecules, which could have an effect on specificity and selectivity.[55]

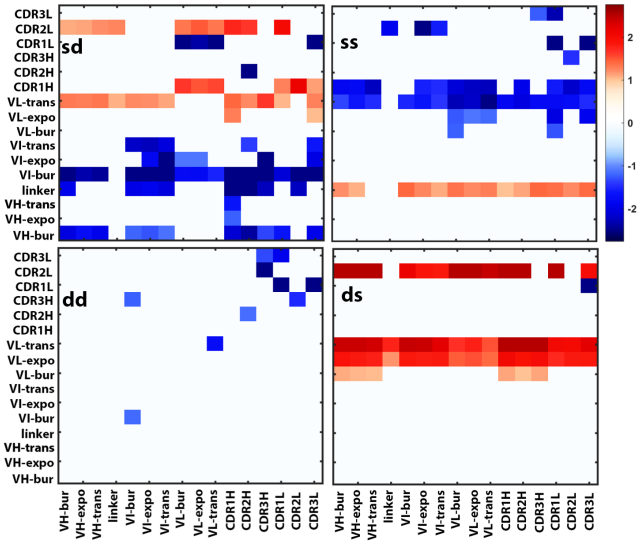


Figure 8: Coarse-grained perturbation response matrix. Each block represents a targeted perturbation (y-axis) and targeted response (x-axis). Within a block the ratio of a targeted response to the mean response over the entire protein is shown on a $\log_2$ scale. White blocks have no signal or insufficient statistical confidence.

Different Perturbation Types: The triad, star and ball perturbation types (defined in Methods) are tested to mimic how different ligands may induce different dynamic responses depending on the nature/type of a perturbation. At the very least, a comparison of the different perturbation types allows sensitivity of the model to be explored. Response and perturbation profiles for each type of perturbation for ss, sd, dd, and ds cases are shown in **Figure 9**. The star and ball perturbation types induce similar response patterns, but with different magnitudes, while the triad perturbation exhibits a distinctly different pattern.

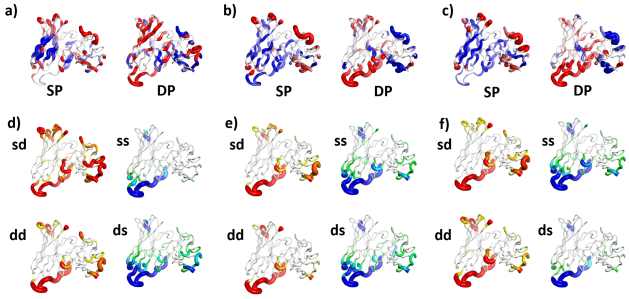


Figure 9: Mean response to different types of stabilizing (SP) and destabilizing (DP) perturbation for a) triad b) star and c) ball.

Linear correlation coefficients are greater than 0.8 when two mean response profiles from triad, star and ball perturbations are

compared across all scFv fragments if the linker is excluded. The linear correlation coefficients between star and ball are about 0.7 to 0.8, whereas between triad and star or ball are below 0.5 always. The linear correlation coefficients fall within these ranges for the ss, sd, dd and ds cases. These results suggest the response profile by scFv antibody fragments to different ligands is highly conserved. Conversely, the propensity of triggering a response depends on the nature of the perturbation and location of the perturbing site. As such, the coarse-grained PRM shown in **Figure 8** changes between triad and ball perturbation types. This means different computational models will lead to different predictions due to the degree of sensitivity in the type of perturbation. Going forward, models that aim to predict allosteric pathways should place more attention on the chemical nature of the perturbation, which can have stabilizing and destabilizing effects together.

Conclusions

Mean destabilizing response in mobility across the native and mutant scFv fragments were markedly conserved for the four combinations of stabilizing or destabilizing perturbations and stabilizing or destabilizing responses. Response profiles from each corresponding CDR-loop were similar across the scFv fragments. These results are interpreted as being a consequence that all scFv antibody fragments studied are functional. However, perturbation locations that induce similar responses are variable. If exposed residues are likely to be destabilized due to solute protein interactions, then this is likely going to lead to destabilizing effects in the VH domain but stabilizing effects in the VL domain. Generally, the linker transfers allosteric signals from the VH domain to the VL domain and vice versa. Decomposition of perturbation and response show that functionally important loops have high mobility in various mutants.

Identifying dynamical couplings help elucidate which residues to target with a certain type of perturbation to increase stability of a protein while maintaining or modifying function. The method employed here decomposes dynamical couplings into stabilizing and destabilizing effects, which provides a powerful approach for understanding sensitivity in structure/function relationships, and in studying dynamic allostery. The method has the plasticity and speed of ENM approaches during perturbation scans, and its accuracy is inherited from the detailed information about atomic fluctuations derived from molecular dynamics simulations.

ASSOCIATED CONTENT

Supporting Information

Supporting Information is available free of charge on the ACS Publications website comprising 20 figures and 2 tables (PDF).

AUTHOR INFORMATION

Corresponding Author

* To whom correspondence should be addressed. E-mail: djacobs1@uncc.edu

Acknowledgments

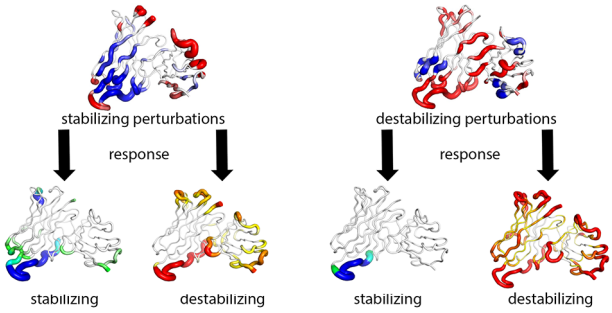
Support for Dr. Azhagiya Singam Ettayapuram Ramaprasad was provided by Medimmune Inc and partial support to DJJ through NIH grant GM101570.

REFERENCES

1. Winter, G.; Milstein, C., *Nature* **1991,** *349*, 293-9.
2. Li, X.; Stuckert, P.; Bosch, I.; Marks, J. D.; Marasco, W. A., *Cancer Gene Ther* **2001,** *8*, 555-65.
3. Corson, T. W.; Aberle, N.; Crews, C. M., *ACS Chem Biol* **2008,** *3*, 677-692.
4. Ahmad, Z. A.; Yeap, S. K.; Ali, A. M.; Ho, W. Y.; Alitheen, N. B.; Hamid, M., *Clin Dev Immunol* **2012,** *2012*, 980250.
5. Lehmann, A.; Wixted, J. H.; Shapovalov, M. V.; Roder, H.; Dunbrack, R. L., Jr.; Robinson, M. K., *MAbs* **2015,** *7*, 1058-71.
6. Miller, B. R.; Demarest, S. J.; Lugovskoy, A.; Huang, F.; Wu, X.; Snyder, W. B.; Croner, L. J.; Wang, N.; Amatucci, A.; Michaelson, J. S.; Glaser, S. M., *Protein Eng Des Sel* **2010,** *23*, 549-57.
7. Spiess, C.; Zhai, Q.; Carter, P. J., *Mol Immunol* **2015,** *67* (2 Pt A), 95-106.
8. Chames, P.; Baty, D., *MAbs* **2009,** *1*, 539-47.
9. Li, T.; Tracka, M. B.; Uddin, S.; Casas-Finet, J.; Jacobs, D. J.; Livesay, D. R., *PLoS One* **2014,** *9*, e92870.
10. Pandini, A.; Fornili, A.; Fraternali, F.; Kleinjung, J., *FASEB J* **2012,** *26*, 868-81.
11. Tompa, P., *Chem Rev* **2014,** *114*, 6715-32.
12. Changeux, J. P., *Protein Sci* **2011,** *20*, 1119-24.
13. Monod, J.; Wyman, J.; Changeux, J. P., *J Mol Biol* **1965,** *12*, 88-118.
14. Cui, Q.; Karplus, M., *Protein Sci* **2008,** *17*, 1295-307.
15. Guo, J.; Zhou, H. X., *Chem Rev* **2016,** *116*, 6503-15.
16. Dokholyan, N. V., *Chem Rev* **2016,** *116*, 6463-87.
17. Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J., *Nature* **2014,** *508*, 331-9.
18. Cooper, A.; Dryden, D. T., *Eur Biophys J* **1984,** *11*, 103-9.
19. Tzeng, S. R.; Kalodimos, C. G., *Nature* **2009,** *462*, 368-72.
20. Lee, A. L., *Biophys Rev* **2015,** *7*, 217-226.
21. Lin, M. M., *J Am Chem Soc* **2016,** *138*, 5036-43.
22. Grutsch, S.; Bruschweiler, S.; Tollinger, M., *PLoS Comput Biol* **2016,** *12*, e1004620.
23. Law, A. B.; Sapienza, P. J.; Zhang, J.; Zuo, X.; Petit, C. M., *J Am Chem Soc* **2017,** *139*, 3599-3602.
24. Gunasekaran, K.; Ma, B.; Nussinov, R., *Proteins* **2004,** *57*, 433-43.
25. Bahar, I.; Chennubhotla, C.; Tobi, D., *Curr Opin Struct Biol* **2007,** *17*, 633-40.
26. Srivastava, A.; Tracka, M. B.; Uddin, S.; Casas-Finet, J.; Livesay, D. R.; Jacobs, D. J., *Biophys J* **2016,** *110*, 1933-42.
27. Tang, Q.; Fenton, A. W., *Hum Mutat* **2017,** *38*, 1132–1143.
28. Atilgan, C.; Gerek, Z. N.; Ozkan, S. B.; Atilgan, A. R., *Biophys J* **2010,** *99*, 933-43.
29. Atilgan, C.; Atilgan, A. R., *PLoS Comput Biol* **2009,** *5*, e1000544.
30. Kumar, A.; Glembo, T. J.; Ozkan, S. B., *Biophys J* **2015,** *109*, 1273-81.
31. Gerek, Z. N.; Ozkan, S. B., *PLoS Comput Biol* **2011,** *7*, e1002154.
32. Zheng, W.; Thirumalai, D., *Biophys J* **2009,** *96*, 2128-37.
33. Zheng, W.; Brooks, B. R.; Thirumalai, D., *Proc Natl Acad Sci U S A* **2006,** *103*, 7664-9.
34. Zheng, W.; Brooks, B. R.; Doniach, S.; Thirumalai, D., *Structure* **2005,** *13*, 565-77.
35. Ming, D.; Wall, M. E., *Phys Rev Lett* **2005,** *95*, 198103.
36. Kalescky, R.; Liu, J.; Tao, P., *J Phys Chem A* **2015,** *119*, 1689-700.
37. Kalescky, R.; Zhou, H.; Liu, J.; Tao, P., *PLoS Comput Biol* **2016,** *12*, e1004893.
38. Yao, X. Q.; Skjaerven, L.; Grant, B. J., *J Phys Chem B* **2016,** *120*, 8276-88.
39. Peng, J.; Zhang, Z., *J Chem Theory Comput* **2014,** *10*, 3449-58.
40. Townsend, P. D.; Rodgers, T. L.; Pohl, E.; Wilson, M. R.; McLeish, T. C.; Cann, M. J., *Biophys Rev* **2015,** *7*, 175-182.
41. Bahar, I.; Lezon, T. R.; Bakan, A.; Shrivastava, I. H., *Chem Rev* **2010,** *110*, 1463-97.

42. Bahar, I.; Lezon, T. R.; Yang, L. W.; Eyal, E., *Annu Rev Biophys* **2010,** *39*, 23-42.
43. Lipkowitz, K. B.; Boyd, D. B., Reviews in computational chemistry. VCH: New York, N.Y., 1998.
44. Zheng, S.; Moores, S.; Jarantow, S.; Pardinas, J.; Chiu, M.; Zhou, H.; Wang, W., *MAbs* **2016,** *8*, 551-61.
45. Sengers, B. G.; McGinty, S.; Nouri, F. Z.; Argungu, M.; Hawkins, E.; Hadji, A.; Weber, A.; Taylor, A.; Sepp, A., *MAbs* **2016,** *8*, 905-15.
46. Kellmann, S. J.; Dubel, S.; Thie, H., *MAbs* **2017,** *9*, 404-418.
47. Papaleo, E.; Saladino, G.; Lambrughi, M.; Lindorff-Larsen, K.; Gervasio, F. L.; Nussinov, R., *Chem Rev* **2016,** *116*, 6391-423.
48. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., *Bioinformatics* **2013,** *29*, 845-54.
49. David, C. C.; Singam, E. R. A.; Jacobs, D. J., *BMC Bioinformatics* **2017,** *18*, 271.
50. Reddy Chichili, V. P.; Kumar, V.; Sivaraman, J., *Protein Sci* **2013,** *2*, 153-67.
51. Cheng, J.; Dul, Q.; Zhang; Ren, Y.; Zhang, B.; Feng, X., *Eur J Gynaecol Oncol* **2016,** *37*, 171-7.
52. Luque, I.; Freire, E., *Proteins* **2000,** *4*, 63-71.
53. Stanfield, R. L.; Takimoto-Kamimura, M.; Rini, J. M.; Profy, A. T.; Wilson, I. A., *Structure* **1993,** *1*, 83-93.
54. Roguin, L. P.; Retegui, L. A., *Scand J Immunol* **2003,** *58*, 387-94.
55. Cantor, R. S., *Biophys J* **1999,** *77*, 2643-7.

TOC


stabilizing perturbations
response
stabilizing
destabilizing
destabilizing perturbations
response
stabilizing
destabilizing