Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 20, Issue Number 2, (2002)

Flexibility of BIV TAR-Tat: Models of Peptide Binding

http://www.jbsdonline.com

Mark Hsieh[1]
Elaine D. Collins
Thomas Blomquist
Brooke Lustig*

Department of Chemistry
San Jose State University
San Jose CA, 95192-0101

[1]Present address:
Abmaxis Inc.
453 Ravendale Dr.
Suite B
Mountain View, CA 94043

*Phone: (408) 924-4968
Fax (408) 924-4945
Email: blustig@science.sjsu.edu

Abstract

A new approach in determining local residue flexibility from base-amino acid contact frequencies is applied to the twelve million lattice chains modeling BIV Tat peptide binding to TAR RNA fragment. Many of the resulting key features in flexibility correspond to RMSD calculations derived from a set of five NMR derived structures (X. Ye, R. A. Kumar, and D. J. Patel, Protein Data Bank: Database of three-dimensional structures determined from NMR (1996)) and binding studies of mutants (L. Chen and A. D. Frankel, *Proc. Natl. Acad. Sci. USA 92*, 5077-5081 (1995)). The lattice and RMSD calculations facilitate the identification of peptide hinge regions that can best utilize the introduction of Gly or other flexible residues. This approach for identifying potential sites amenable to substitution of more flexible residues to enhance peptide binding to RNA targets could be a useful design tool.

Introduction

The interaction of RNA with proteins is critical to many important biological processes such as transcription, translation and regulation (1, 2). There are several classes of proteins that can interact with RNA. One class, the Arg-rich motif (3), generally does not appear to have a distinct secondary structure until it binds to the RNA (4). The interaction with the RNA scaffold appears to constrain the binding region into the optimal structure for interaction.

The *trans*-activating proteins (Tat) for Bovine Immunodeficiency Virus (BIV) and Human Immunodeficiency Virus (HIV) are considered Arg-rich and bind to a target RNA terminal activating region (TAR) to initiate transcription. Tat peptides can be used to model these interactions. Recent NMR solution studies of BIV TAR-Tat binding species by Frankel and coworkers (5) and Patel and coworkers (6) show specific contacts between the TAR and Tat in the major groove. The TAR fragment 28nt was the same in both studies, but the Frankel group's 14-residue Tat peptide is three amino acids shorter in size. In a somewhat different approach, the Patel study uses separately determined NMR structures for Tat peptide and TAR fragment. The modeling then best fits together the two elements. A strong similarity in overall binding species structure is indicated by both groups (4).

In vitro and *in vivo* experiments show that specific base pairs as well as bulges are involved in the binding in BIV (7-9). The best-defined RNA-peptide interactions span the eleven residues Arg70 through Arg80. These define a peptide β-hairpin wedged into the region of the two RNA helical stems that are stacked on one another. There is a correlation between the flexibility afforded by the bulges of unbound BIV TAR and their affinities for Tat (10), suggesting that RNA bulge flexibility may play a role in binding between RNA and peptide. Furthermore, there are a number of critical glycine residues in the Tat peptide (4), suggesting some key role for them in binding.

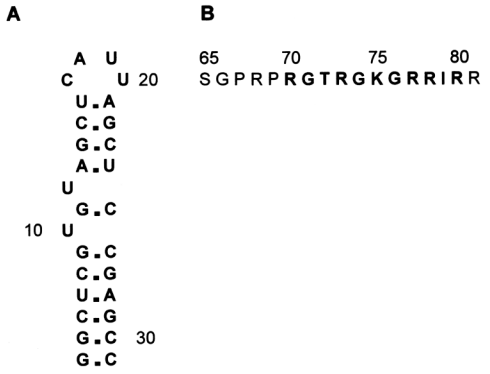

Figure 1: BIV TAR and Tat. A) Secondary structure of RNA BIV TAR fragment. B) Wild-type sequence of Tat binding peptide. Bold key indicates binding portion used in lattice modeling.

There appears to be a significant role for structural flexibility in protein binding and related functions (11, 12). Wells and coworkers have used phage display to optimize peptide binding to Fc fragment of immunoglobulin (13). There is evidence that two methionine residues form a pocket that facilitates access by the peptide, increasing the range of peptide states (i.e. increasing its flexibility), hence enhancing its ability to bind to the protein. Also, there is evidence that additional flexibility is afforded by the hinge region adjacent to these sites (11). Such hinge flexibility allows the accommodation of mutations and the binding of potentially different ligand structures and species. This also has been noted for a variety of cytokine ligand-receptor complexes (14). Similarly, enzyme function for engineered lactose permease is restored upon the addition of Gly substitutions at key interface sites between helices (15). It is also of recent interest that the addition of a Gly to the active site of kinases increases the binding of an inhibitor (16).

Lattice calculations have used rigid double-helical stems connected by a bulge region to model HIV and BIV TAR RNA. These lattice calculations showed that increased RNA bulge size increased access for Tat binding, suggesting that flexibility improved binding (10). This helped to explain earlier experimental results (8, 17) that showed increased binding occurred with larger, i.e. more flexible, RNA bulges. This suggests that increasing flexibility in the peptide could also increase binding of peptide to RNA.

Schneider and coworkers have characterized control site regions of known DNA-protein contacts (18) using an expression derived from information theory for the Shannon entropy. Estimates for the relative flexibility for BIV TAR bulge mutants have already been determined from direct calculations of molecular entropy related to the relative number of generated lattice chains (10). Other experimentally consistent features of wild-type and mutant hammerhead ribozyme flexibility have also been characterized by such methods (19).

Here flexibility is defined in the simplest structural sense where the least constrained part of the structure more readily undergoes conformational rearrangements. The increased local flexibility is indicated by a larger configurational entropy. Such flexibility appears associated with the ability of local structure to accommodate mutation at that position. We will explore the utility of calculating relative flexibility of key BIV Tat peptide residues from structure and sequence.

Methods and Results

RMSD Calculations for BIV TAR-Tat

An ensemble of five NMR structures determined by Patel and coworkers (6, 20) are used to calculate RMSD (root mean square deviation) values for the 28-nt BIV TAR (Figure 1A) binding the 17-mer Tat peptide (Figure 1B). Here we calculate RMSD values for the backbone O3', the sugar O4' and purine N1or pyrimidine N3 atoms (Figure 2) of the 28-nt BIV TAR fragment and for the Cα, Cβ and Cγ atoms (Figure 3) of the 17-mer Tat peptide. The RNA backbone (Figure 1A) appears fairly constrained, except at the ends of the two A-form stems, one of which (TAR 17-20) describes the hairpin loop. This suggests that there is limited stem movement. Some fluctuation (Figure 2) is also noted with respect to the bulge regions (TAR 10-12). Interestingly, base fluctuations involving N1 (purine) or N3 (pyrimidine) atoms show very significant variation with respect to O3' backbone and O4' sugar atoms at bulge position 12 and the hairpin loop (17-20) region. For the BIV Tat peptide there are three regions of flexibility (Figure 3) that correspond to the interior of the binding domain (Tat 70-80) and the N and C terminal regions. In the interior domain one can describe a flexible hinge region that includes residues Lys75 and Gly76 and as well as at Arg78.

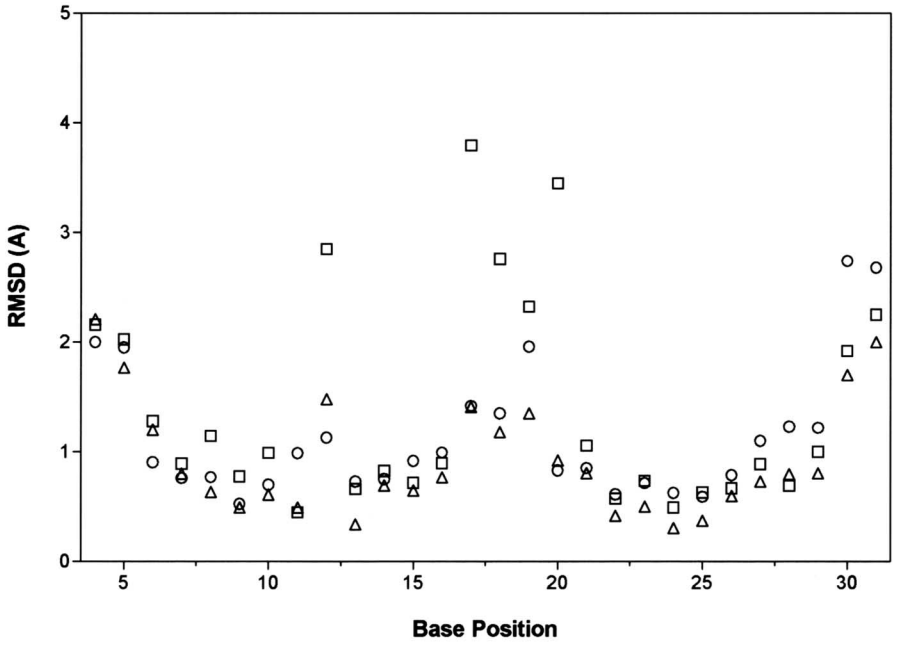

Figure 2: Fluctuations (A) indicated by RMSD (root mean square deviation) for selected RNA atoms of BIV TAR RNA fragment binding to Tat peptide, determined from ensemble of five structures determined from NMR (20). Circles indicate RMSD values for O3' atoms, triangles for O4' and squares for N1(purine) or N3 (pyrimidine) atoms.

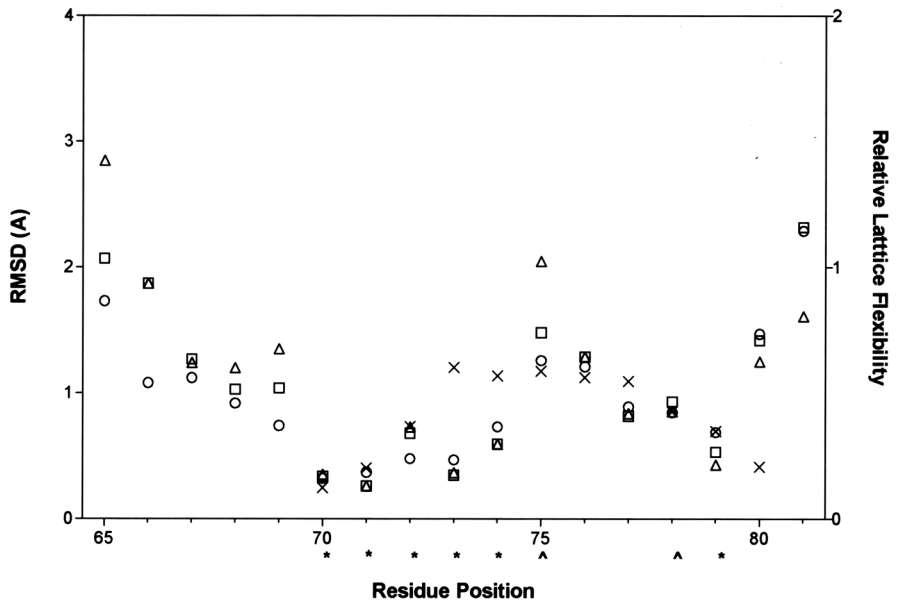

Figure 3: Measures of flexibility for of BIV Tat peptide binding to TAR fragment. Fluctuations (A) indicated by RMSD (left axis) for selected peptide atoms, determined from ensemble of five structures determined from NMR (20). Stars on abscissa indicate consensus RNA base-amino acid contacts (5, 6), and carats note positions Lys75 and Arg78 chosen for possible substitution by Gly. Open circles indicate RMSD values for Cα atoms, squares for Cβ and triangles for Cγ atoms (left axis). The crosses correspond to relative flexibility (right axis) of a residue calculated from a complete generation of all model chain structures on a lattice.

Sequence Variability, Entropy and Flexibility: BIV Tat Mutability

For peptide or protein sequence an expression for sequence entropy S_k at some amino acid position k can be expressed as

$$S_k = -\sum_{j=1,20} P_{jk} \ln P_{jk} \qquad [1]$$

where the probability Pjk at some amino acid sequence position k is derived from the frequency f_{jk} for an amino acid type j (e.g. Lys) at sequence position k for N aligned or otherwise correlated sequences. One would expect that in general the larger the configurational entropy associated with an amino acid position in the binding region of the Tat peptide the greater the local flexibility. It is a reasonable hypothesis that sequence entropy often correlates with the configurational entropy.

Direct calculations of sequence entropy from sequence variability for just BIV Tat is not possible because of the very limited number of sequences. However, Frankel and coworkers (7) classify in the BIV Tat region 70-80 a set of highly mutable

sequences, involving changes in residues Lys75, Arg78 and Arg80, that still result in binding to TAR fragment. We can estimate an upper boundary of 3 for these three residues using eq. [1] if we assume that the term Pjk is equally partitioned as 1/20 for all twenty amino acid types. A lower boundary of zero is indicated for those terms that remain highly conserved. An intermediate value suggests some 20% possible variants, suggesting that even intermediate entropy values can be descriptive of what are in practical terms still very variable sequences. The RMSD plot for BIV Tat (Figure 3) shows that within the binding domain 70-80 the most flexible residues include Lys75 and Arg80.

The expression for sequence entropy in eq. [1] does indicate some sort of partitioning. An increased mutability at a residue position corresponds to more terms in eq. [1], resulting in increased sequence entropy. Now we can also reasonably assume that the greater the flexibility at a residue position the more able it is to accommodate mutation. This suggests a correlation between sequence entropy and local flexibility. The assumption is that the most flexible possible RNA-peptide species would have all its amino acids equally capable of contact with each nucleotide. This is suggestive of a peptide sequence position that is totally mutable. Now we can calculate for each possible base-amino acid some square of the magnitude in the difference of the number of such contacts f_{jk} with respect to the average frequency $\bar{f}$ for all contacts, such that

$$(f_{jk} - \bar{f})^2 \sim 0 \qquad [2]$$

We assume in this case of maximum flexibility that each f_{jk} value is close in number to $\bar{f}$, the average number of contacts. If we sum these squares of the differences from the mean with respect to each of the j nucleotides we can write a proposed expression for relative flexibility F_k

$$F_k = 1/ (\sum (f_{jk} - \bar{f})/N)^2)^{1/2} \qquad [3]$$

where N is the total number of possible contacts, the same as the number of lattice chains. The relative flexibility F_k at some amino acid position k would be very large as we would expect, given the condition of maximum flexibility noted in eq. [2].

However, a more rigid system will favor certain nucleotide-amino acid contacts relative to others, so for a given amino acid i we could expect

$$(f_{jk} - \bar{f})^2 > 1$$

where according to eq. [3], we would calculate, as expected, small values for relative flexibility F_k.

Lattice Modeling of BIV-Tat

The lattice used in the calculation is a modified simple cubic lattice so that a point centered at the junction of eight cubes would have six edge and twelve face diagonal moves in space. The moves here are specifically associated with the peptide conformation attached to the two fixed stems for RNA. The lattice spacing used was 6. 15A because it is well suited as a characteristic distance for the spacing between O3' atoms of a canonical A-form RNA (21). The same lattice is applied to the peptide chain generation using two amino acids per lattice move (22), to achieve proper scaling between the protein and nucleic acid units, as well as to allow complete enumeration.

Figure 1A shows the bulge region of the 28nt (nucleotide) TAR RNA fragment. This bulge region between the stems can be viewed as a 3nt bulge opposite a single nucleotide, but with that single nucleotide paired to its opposite. This results in

two 1nt bulges. We have used an 11-residue peptide corresponding to wild-type (WT) Tat sequence Arg70 to Arg80 (Figure 1B). Odd numbered peptide positions are assigned to the midpoints of each chain segment. All possible RNA-peptide lattice structures are generated, consistent with two fixed stems and Arg70-Gly71 affixed to G14 and Ile79-Arg80 affixed to U10, and are consistent with NMR determined three-dimensional structure (5, 6) and restrictions in lattice positioning.

In order to reduce computation time, each of the two six-base pair RNA stem regions of BIV TAR are assigned fixed coordinates. Each of the two helices is constructed by assignment of twelve lattice points to a superimposed six-base pair A-form RNA, using the framework of O3' atoms (see Figure 4). Two of the assigned points are then rearranged to close potential holes that could allow for a lattice chain to pass through. The two six-base pair lattice stem structures are arranged with the second rotated 180° about the helix axis and translated four lattice units along the helix axis relative to the first. The two stems are fixed in position at a distance consistent with stem-stem stacking. Complete enumeration of the bulge region and the peptide required at least 200 minutes CPU time on a Silicon Graphics O2 workstation.

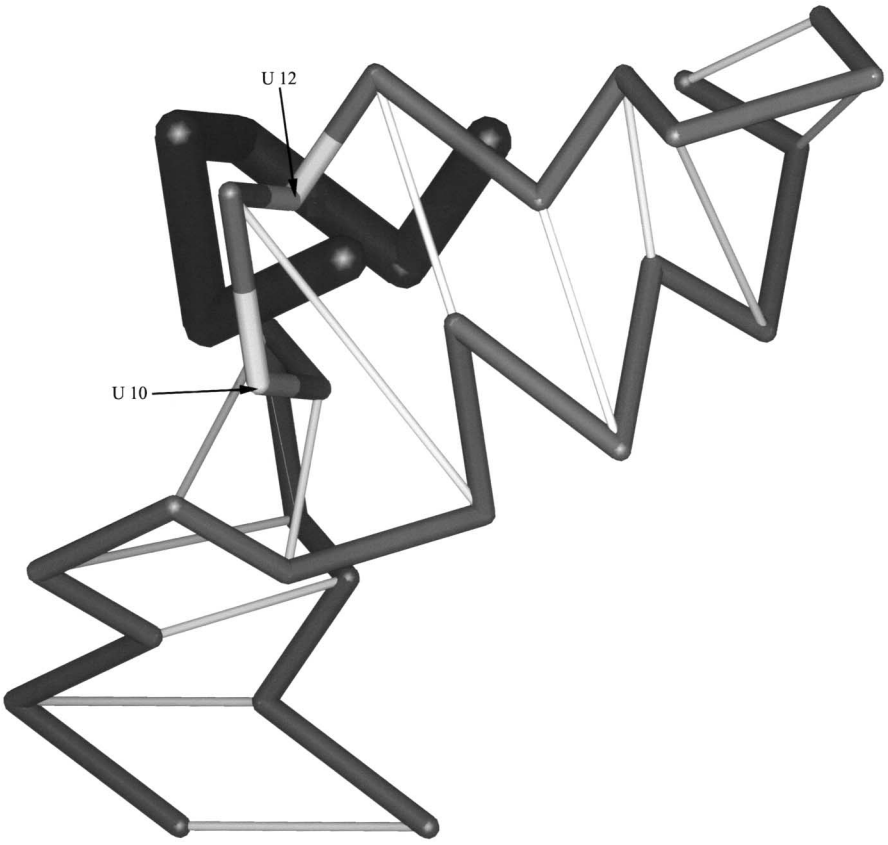

Figure 4: Sample Lattice Structure for BIV TAR-Tat. Base-pairing indicated by thin lines. Note Tat peptide has five linkers (2 residues per linker) shaded black and the two annotated linkages are bulge nucleotides U10 and U12.

Calculating Flexibility from Lattice Calculations

We propose calculating flexibility for that RNA-peptide complex using a reduced representation of the lattice chain calculations involving the contact map shown in Table I. Here each cell of the table represents f_{jk}, the number of contacts for base at position j (rows Table 1) with amino acid at position k (columns). From eq. [3] we can calculate for each amino acid position k a relative flexibility F_k, where the average frequency of contact $\bar{f}$ is 1,254,266 and out of a total of 12,047,142 possible single contacts.

Lattice Based Energy Calculations

From the complete enumeration of the some twelve million lattice chains modeling BIV TAR-Tat we have calculated the energies for wild-type and a set of eleven BIV

Tat peptides and TAR fragments. Existing RNA base-amino acid interaction potentials (23) were utilized. The potentials are divided into major groove and non-major groove values, reflecting the different types of interactions possible with polynucleotides. However, in this case, only major groove potentials were applied because contacts were effectively confined to the major groove. The algorithm calculates the energy of each lattice chain, substituting the appropriate base or amino acid type at their sequence position in all lattice structures. A contact between any given amino acid and base is defined as being within a range of 12. 30 thru 24.60A. Subsequently, native as well as mutant RNA and peptide sequences are superimposed on each lattice structure. The energy for each lattice structure is a sum of all base-amino acid energies.

Table I

Contact frequencies for wild-type BIV RNA bases with Tat peptide amino acids.

	Amino Acid										
	R70	G71	T72	R73	G74	K75	G76	R77	R78	I79	R80
Base											
4g	0	0	0	0	0	936	13020	4648	34190	7951	24943
5g	0	0	0	0	18475	0	27475	127	43008	467	102132
6c	0	0	408391	82181	185953	72596	188699	106528	304167	2109889	577279
7u	0	0	0	29013	352326	124046	369133	161319	472660	275167	651176
8c	0	401976	1334134	893002	1218668	804733	1350275	1018560	1921696	1998176	3605176
9g	3958901	1044531	1044531	510380	922879	542345	1103245	803151	1848056	1966210	4430234
10u	4841827	2473002	3037246	1970765	2654142	2058364	3067330	3039250	5132284	6971643	12047142
11g	2638291	1827796	2482779	1803897	2383589	1823324	2631255	2346509	3619567	4010599	6170110
12u	4152073	2367308	2776624	1914827	2446411	1665799	2457456	1934355	3035733	2979518	4565580
13a	9126149	9126149	4436176	2205943	3322798	2083777	2767675	1952109	3000776	3044111	4382679
14g	2047142	12047142	5079686	1190865	2084554	483159	1262924	382059	1238637	545624	1823079
15c	12047142	1983082	1983082	766214	645618	227228	360236	135527	248111	76011	186565
16u	2920993	3150566	2560328	757585	1023235	304890	471948	129429	245458	21359	2315
17c	0	0	32279	2206	9176	405	1136	0	0	0	0
18a	0	0	0	0	49513	6561	21017	0	0	0	0
19u	3958901	2624767	2624767	1357624	1092147	519472	582044	321776	444405	240865	405069
20u	0	0	713261	253209	576583	277380	415203	163890	285122	90826	125806
21a	0	0	479843	260530	476821	265273	358865	180647	298599	161500	218656
22g	12047142	1799697	1799697	772161	731291	246621	409939	158001	417239	225660	679461
23c	2920993	1589229	2614477	832730	1095543	463260	790385	351596	758851	270333	619769
24u	12047142	4635433	4635433	2352549	2329734	1401294	1985898	1424201	2432638	2620208	4256875
25c	1638067	737159	1426981	835937	1381053	966316	1503394	1207561	1951472	1681129	2758306
26c	0	0	0	0	261175	69817	379247	119838	624389	332703	1239751
27g	0	0	549183	329247	921629	676122	1434991	1122755	2208326	2054826	3514063
28a	0	0	0	0	57479	20335	172357	78679	397692	209089	863619
29g	0	0	0	0	0	0	8459	1109	42075	10810	123993
30c	0	0	0	0	0	0	0	92	1292	283	13095
31c	0	0	0	0	0	0	0	0	119	0	0

The individual amino acid energies are themselves a weighted sum of the energies of that particular amino acid and all adjacent bases. The weights are set to 1.0, 0.5 and 0.25 for two, three and four lattice unit separations (6.15A per lattice unit), respectively. This convention allows for a well-defined selection of inclusive points compared to just using radial distances in lattice space. The base-amino acid pairs that are within this lattice distance range are included in the energy summation. No attempt is made to selectively identify the one dominant base-amino acid contact.

The calculations from lattice for the relative amino acid flexibility of BIV TAR-Tat show a fair amount of agreement with the NMR experimentally derived RMSD values as shown in Figure 3. In addition less lattice flexibility is indicated at Tat position 80 (Arg for WT) than indicated directly by NMR. This latter anomalous behavior is best explained as an artifact of fixing Arg80 relative to U10 of the TAR fragment. This needed procedural constraint is reasonable given that the actual Ile79-U10 contact (5) is itself effectively constrained (see Table I) in the first lattice move (two residues per lattice move) from that end of the attached peptide chain. Also, Tat position 73 (Arg for WT) shows more lattice flexibility than indicated directly by NMR. Here calculations of lattice flexibility are not constrained implicitly or explicitly by the energetics associated with a contact between a particular residue, like Arg73, with various bases.

There was significant success in checking for the likelihood of strong stabilization due to a substitution. We assume that the magnitude for such stability is greater than kT, 0.6 kcal/mol. This level would be a key discriminator in assessing the suitability of possible substitutions. Average energies were characterized for WT and 11 representative mutations (Arg70, 73 and 77 to Lys, Thr72, Gly74 and Ile79 to Ala; U10 to A or C, G11 to C, U12 to C and C25 to U) among all lattice chain conformers. These mutants were all experimentally determined by Chen and Frankel to bind with lower affinity than WT (7, 8). In our energy calculations involving 12 million lattice chains, two of six peptide variants (Gly74Ala, Ile79Ala) and none of five TAR RNA mutants falsely indicated strong stabilization. No significant differences in the number of these sorts of false positives were obtained for the 10% of chains with the lowest energies. Most importantly, those false positives identified from the complete and low-energy sets of lattice configurations exclude mutations involving Arg, which has one of the best characterized and most stabilizing base-amino acid potential energies (23). All the complexes involving mutations of Arg at BIV Tat positions 70, 73 and 77 were calculated to be less stable.

Discussion

Consensus BIV TAR-Tat NMR Structural Features

There are 3 major consensus contacts (4), Arg70-G14, Arg73-G11 and Ile79-U10, for the BIV Tat binding domain 70-80 as determined by NMR (5, 6). All three include RNA base-amino acid contacts. In addition there are contacts indicated for Thr72 and Arg77, but the specifics vary significantly for Thr22 (4). The position of Arg77 is not well defined in either NMR structure. Overall there is great similarity in the topology of folding, including the formation of a β hairpin deep in the major groove, where Gly71 and Gly74 are critical for that turn.

There are apparent differences in local flexibility for the two NMR-determined sets of three-dimensional structure (5, 6), as measured by average base fluctuation per residue calculated from independent molecular dynamics (MD) calculations by Kollman and coworkers (24). The Puglisi-derived MD calculations are anomalous with respect to the various flexibility profiles derived by sequence and lattice as opposed to the general pattern of agreement with the Patel-derived MD data. The most glaring anomalous feature for the Puglisi-derived data involves the apparent high-degree of fluctuation calculated for Arg70. Further analysis of the Puglisi-derived trajectories by Kollman and coworkers (24) shows Arg70 pulling away from the RNA. This appears inconsistent with the optimized NMR structure actually determined by Puglisi, which clearly suggests hydrogen bonding of Arg70 with the O6 and N7 of G14 (5).

One possible explanation for the anomalous MD calculation (24) is that an average structure was used as the starting structure (5) unlike the Patel-derived MD calculations (6). This might preclude starting at some local minimum of the energy surface that would correspond to one of the structures indicated in the original ensemble of twenty. Overlays of the two sets of MD determined fluctuations are generally consistent with the RMSD results directly calculated from the one ensemble of five structures (20). Missing NOEs (nuclear Overhauser effects) are problematic with any single set of NMR data (25, 26), so overlays are useful in filling any gaps.

BIV Tat Flexibility: Structure and Sequence

Extensive quantitative measures of flexibility scaled from sequence data are not currently available because of the limited number of variant BIV Tat peptides. However, it is quite interesting that flexibility indicated by peptide mutant binding studies qualitatively matches NMR derived fluctuation data for the BIV TAR-Tat binding. Gel-shift and in vivo activity assays show minimal effect at Lys75, Arg78 and Arg80 due to various mutations of the binding domain at position 70 thru 80

for the 17-residue BIV Tat (7, 8). Similarly, key residues sensitive to mutation include Arg70, Arg73 and Ile79, which are unambiguously determined to be residues involved in contact, and structurally key residues Gly74 and Gly76. We can assume that at least these two Gly residues are by themselves indicative of significant intrinsic structural flexibility. NMR derived RMSD values ranging from Lys75 thru Ile78 indicate the most significant flexibility. Only the conserved nature of Arg77 as indicated by the mutational studies seems inconsistent with its significant structural flexibility, as measured by RMSD.

One possible interaction for Arg77 is with the phosphate backbone (5). We have shown that DNA and RNA interactions not specifically involving base-amino acid contacts do not show strong sequence dependence (23). However it is clear that the bivalent cationic character of Arg side-chain is most suitable for stabilizing interaction with phosphate backbone (27). This dichotomy can be reconciled noting RNA phosphate-peptide interactions positions, dictated by a matrix of base-amino acid interactions, involve Arg interaction with the phosphate that includes some local structural flexibility. Similarly, Patel and coworkers indicate the position of the Arg77 side-chain is variable, and may include a variety of possible interactions with the G9-C26 nucleotide pair (6).

Flexibility Role for Gly

There are a number of critical glycine residues in the Tat peptide, suggesting a role for peptide flexibility in this system. Tat peptide variants involving substitutions of Gly at positions 71, 74 and 76 with less flexible Ala showed decreased binding (7). This is consistent with the Gly74 and Gly76 better facilitating the β-hairpin indicated between Arg73 and Arg77 (4). Possible candidates for substitution to Gly to enhance flexibility at the ends of possible hinge-like regions include Lys75 and Arg78. These should allow the relevant amino acids, including Arg70, Ile79 and Arg73, better access to their respective RNA contacts flanking and in the bulge region.

Conclusion

The dominant features of flexibility for BIV TAR-Tat can be successfully calculated from the base-amino acid contact maps of exhaustive lattice explorations. Other details can be deduced from additional energy calculations of the lattice chains including the identification of Arg73 as another contact residue in contrast to Arg77, which is not unambiguously a contact residue. The identification is clear given the relative position of Arg73 with respect to its neighbors' calculated lattice flexibilities. Residues like Arg73 are unlikely to be as flexible as indicated by calculations using just the lattice contact maps that are dominated by excluded volume considerations. More extensive analysis of energies and their weightings await more refined nucleotide-amino acid energy potentials.

The expression for calculating flexibility from contacts may be more universal in its application, for example it could be applied to other sets and types of contact maps of biomolecular interactions. One other use is to help describe the apparent connection between sequence variability (mutability) and configurational entropy of a protein. It is also clear that calculations from one set of NMR data for RMSD values and related MD trajectories may require additional independent verification. Some combination of analysis involving lattice and/or NMR, as described here, does allow the visualization of possible hinge-like regions that may indicate those peptide substitutions that enhance binding.

Acknowledgments

We thank Dr. A. Waleh for helpful discussions. Support in part by grants from NIH (AI39372) and Calif. State Univ. Program for Education and Research in Biotechnology (CSUPERB).

1. C. G. Burd and G. Dreyfuss, *Science 265*, 615-620 (1994).
2. A. D. Frankel, *Curr. Opin. Structr. Biol. 10*, 332-340 (2000).
3. I. A. Laird-Offringa and J. G. Belasco, *Nature Structural Biology 5*, 665-668 (1998).
4. N. L. Greenbaum, *Structure 4*, 5-9 (1996).
5. J. D. Puglisi, L. Chen, S. Blanchard, and A. D. Frankel, *Science 270*, 1200-1203 (1995).
6. X. Ye, R. A. Kumar, and D. J. Patel, *Chemistry & Biology 2*, 827-840 (1995).
7. L. Chen and A. D. Frankel, *Proc. Natl. Acad. Sci. USA 92*, 5077-5081 (1995).
8. L. Chen and A. D. Frankel, *Biochemistry 33*, 2708-2715 (1994).
9. C. A. Smith, S. Crotty, Y. Harada, and A. D. Frankel, *Biochemistry 37*, 10808-10814 (1998).
10. B. Lustig, I. Bahar, and R. L. Jernigan, *Nucleic Acids Res. 26*, 5212-5217 (1998).
11. E. J. Sundberg, and R. A. Mariuzza, *Structure 8*, R137-R142 (2000).
12. G. Zaccal, *Science 288*, 1604-1607 (2000).
13. W. L. Delano, M. H. Ultsch, A. M. de Vos, and J. A. Wells, *Science 287*, 1279-1283 (2000).
14. J. A. Wells and A. M. de Vos, *Ann. Rev. Biochem. 65*, 609-634 (1996).
15. A. B. Weinglass, I. N. Smirnova, and R. H. Kaback, *Biochemistry 40*, 769-776 (2001).
16. A. C. Bishop, J. A. Ubersax, D. T. Petsc, D. P. Matheos, N. S. Gray, J. Blethrow, E. Shimizu, J. Z. Tslen, P. G. Schultz, M. D. Rose, J. L. Wood, D. O. Morgan, and K. M. Shokat, *Nature 407*, 395-401 (2000).
17. K. M. Weeks and D. M. Crothers, *Biochemistry 31*, 10281-10287 (1992).
18. P. P. Papp, D. K. Chattorja, and T. D. Schneider, *J. Mol. Biol. 233*, 219-230 (1993).
19. B. Lustig, N. H. Lin, S. M. Smith, R. L. Jernigan, and K.-T. Jeang, *Nulceic Acids Res. 23*, 3531-3538 (1995).
20. X. Ye, R. A. Kumar, and D. J. Patel, Protein Data Bank: Database of three-dimensional structures determined from NMR (1996).
21. B. Lustig, D. G. Covell, and R. L. Jernigan, *J. Biomol. Struct. Dyn. 12*, 145-161 (1994).
22. D. A. Hinds and M. Levitt, *Proc. Natl. Acad. Sci. USA 89*, 2536-2540 (1992).
23. B. Lustig, S. Arora, and R. L. Jernigan, *Nucleic Acids Res. 25*, 2562-2565 (1997).
24. C. M. Reyes, R. Nifosi, A. D. Frankel, and P. A. Kollman, *Biophys. J. 80*, 2833-2842 (2001).
25. D. E. Wemmer and P. G. Williams, *Methods Enz. 239*, 739-767 (1994).
26. H. Zhou, A. Vermeulen, F. M. Jucker, and A. Pardi, *Biopolymers 52*, 168-180 (2000).
27. W. Saenger, "Principles of Nucleic Acid Structure" Springer-Verlag, New York, 387 (1984).

Date Received: June 4, 2002

Communicated by the Editor Valery Ivanov