Journal Title: Protein Eng
ISSN: 0269-2139
Article Author(s): De Filippis V
Article Title: Predicting local structural changes that result from point mutations.
Year: 1994 Oct
Volume: 7
Issue: 10
Pages: 1203-8
PMID: 7855135

Predicting local structural changes that result from point mutations

*V. De Filippis, C. Sander**, G. Vriend**.*

** CRIBI-Biotechnology Centre, via Trieste 75, 35131 Padua, Italy.*
*** EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg, Germany.*

Abstract

Point mutations are frequently used to explore the structure or function of proteins. The ability to predict the structural effects of point mutations would make the planning of such experiments more reliable. We have now derived a set of detailed predictive rules, based on the comparison of crystal structures of point mutant and wild type in 83 cases. Despite the surprising simplicity of these rules, they describe well the conformational changes in 85 percent of all presently available point mutant structures.

Introduction

A single mutation in a protein can dramatically influence its stability (e.g. Dao-pin et al., 1991) or function (e.g. Grenzin et al., 1992). However, large structural changes as a result of one single mutation are rare (e.g. Thunissen et al., 1993). Generally, the overall structure of proteins does not change upon introduction of a point mutation. If a mutated amino acid does not fit well into its environment, conformational adaptations are made mainly by the mutated residue, to a lesser extent by its entire secondary structure element (e.g. Laughton, 1994), and rarely by its three dimensional contact partners.

Prediction of the structure of a mutated protein is a prerequisite for the design of point mutations aimed at alteration of functional characteristics of a protein. Unfortunately, calculation or even estimates of free energy changes are insufficiently accurate to be used as a routine tool for predicting structural changes. On the other hand, empirical methods based on a variety of approaches work quite well for some purposes (e.g. Abagyan and Totrov, in press; Lee and Levitt, 1991; van Gunsteren and Mark, 1992; Simonson and Brunger, 1992). A reliable and general theory of protein conformational changes, however, does not yet exist.

To circumvent the difficulties with energy calculations, we exploit the fact that certain local structural features are repeated many times in different protein structures. Intuitively, frequent occurrence of features in otherwise different proteins implies that these characteristics are beneficial for folding and stability, and/or that they are easily reached in an evolutionary mutation-selection process. By extrapolation, such features can be used for the prediction conformational changes as a result of point mutations.

We have therefore developed a statistical method for predicting structural changes that is based on searches in the database of 3D structure. Given a residue to be mutated, we search the database for residues with a similar environment and then assess which rotamers in the database are statistically preferred and sterically admissible in the current structure. Using simple rules based on this scheme the local structural environment of a point mutation can be predicted correctly in the majority of all cases. The method is based on simple statistical arguments, and does not require time consuming energy calculations.

Methods

A variety of methods in the program WHAT IF (Vriend, 1990) was used to analyse 83 high resolution X-ray mutant structures: 54 T4-lysozymes, 5 human lysozymes, 8 carbonic anhydrases, 2 interleukin-1b's, 2 isocitrate dehydrogenases, 3 chloramphenicol acetyl transferases, 2 a-lytic serine proteases, 1 ferredoxin, 1 staphylococal nuclease, 1 renin, 1 cytochrome P- 450, 1 Ras-p21 GTPase domain and 2 arabinose binding proteins. Tables 1A and 1B summarise the analysed structures. We only analysed mutant structures that were crystallised under similar conditions as the wild type. X->Gly, Gly->X, X->Pro and Pro->X mutations were excluded when considerable backbone changes were observed. Several mutants (e.g. E46Q in RNase T1 (Grenzin et al., 1992) and C20A in ferredoxin (Stout, 1989)) were not used because of an altered mode of binding of a co-factor. Several other mutants were not included because the wild type and mutant structures were determined by NMR, and showed too many small differences (e.g. K47E in hirudin (Folker et al., 1989)).

For every mutant, a position-specific rotamer analysis was performed. This technique has been described in detail elsewhere (Vriend and Eijsink, 1993; Vriend et al., 1994; Jones and Thirup, 1986). Briefly, a rotamer distribution at a certain position is determined by extracting from a database of non-redundant protein structures (Hobohm et al., 1992) all suitable fragments of length 5 or 7 residues (7 in helix and strand, 5 in case of irregular local backbone). Suitable fragments are those that have a local backbone conformation similar to the one around the evaluated position, and have the same residue type at the central position. In these analyses, the RMS deviation of the alpha carbon positions between the structure and the database fragment was kept below 0.5 □ngstrom. The rotamer distributions were then used to answer the following questions: ¥ What are the differences between the rotamer of the residue mutated by protein engineering and the position specific rotamer distribution obtained from the database of natural proteins? ¥ Which set of rules would have allowed us to predict the structure of the mutated residue correctly?

Results and discussion

Several rotamer libraries have been described (e.g. Morris et al, 1992; McGregor et al, 1987; Ponder and Ridchards, 1987; Dunbrack and Karplus, 1994). In most studies the chi1 distribution of the twenty residues was analysed as a function of secondary structure. In some cases the ends of secondary structure elements were treated separately. The position specific rotamer distributions used here allow for different rotamer distributions at each postion and thus provide a more refined description than do standard rotamer libraries. This is especially true near the ends of secondary structure elements (Vriend et al 1994).

X -> Ala and X -> Gly mutation

We have analysed the structural consequences of 17 X -> Ala and 6 X -> Gly mutants where X is a more voluminous residue. In these cases one can be sure that no strain is introduced by the mutation. If major changes in the structure around the mutated residue would be observed then this would be a strong indicator of strain in the environment of the mutated residues. All 24 mutants of this kind found in the June 1993 version of the PDB are listed in table 1B. One rarely sees large rearrangements of the residues in the direct environment. Seldom does chi1 change by more than ten to twenty degrees (except in two cases, PDB files 1L47 and 1LAA, where a glutamic acid rotates by about 160 degrees to make an hydrogen bond). The only three residues that showed significant rearrangements in these 23 cases are methionines. It is tempting to try to derive residue-specific rules, but the number of examples is still far too small for such an analysis to be statistically significant. Recently Buckle et al. (Buckle et al., 1993) made five mutations in the hydrophobic core of barnase. In all five cases a residue was replaced by a similar but smaller one, and they observed that the mutated residue moves to the greatest extent, normally in the direction of the cavity.

Figure 2 needs to be inserted here

Figure 2. Some examples of database derived rotamer distributions superposed on the experimentally determined structure. The experimental rotamer and the backbone of the 5 residue fragment on which it is centred are shown (red=oxygen, blue=nitrogen, yellow=sulphur, green=carbon). The database extracted rotamers are coloured blue to red for best fitting to poorest fitting respectively. Where applicable residues in the direct environment are also displayed.
A: Two rotamers observed. Most populated rotamer selected. Thr 109 -> Asp in lysozyme (PDB file 1L62). This example is representative of the largest number of cases.
B: Best rotamer shows atomic clashes, second best rotamer selected. Val 149 -> Cys in Lysozyme (PDB file 1L53).
C: Rotamers too diffuse. Salt bridge (green dashed line) determines position. Thr 115 -> Glu in Lysozyme (PDB file 1L37).
D: chi1 correct. chi2 rotates away to optimise hydrogen bonding (indicated by green dashed lines). Ser 38 -> Asp in lysozyme (PDB file 1L19). The backbone of the helix capped by this residue is shown, and intra-helix hydrogen bonds are indicated by dashed lines. Five out of eight residues have the same chi1, but four of those have a chi2 angle 30-60 degrees away from the experimentally observed rotamer.
E: Special case. Best rotamer for Thr 157 -> Asp is observed in PDB file 1L04. A rotation of chi2 would lead to an extra hydrogen bond with a Thr two residues away in the sequence. This is indeed observed as an alternative conformation in PDB file 1L05. The Thr and the hydrogen bond are shown too.
F: chi1 rotated ~20 degrees away from the optimal rotamer. Ile 3 -> Thr in Lysozyme (PDB file 1L18).
G: Optimal situation. All database rotamers correspond to the experimentally observed position. Ala 98 -> Val (PDB file 1L48).

In case of the X->Ala and X->Gly mutations little change is observed in the direct environment of the mutated residue. This trend extends to other types of mutations. Almost all big changes introduced by point mutations arise from residues with long side chains (arginine, glutamic acid, etc.) at the protein surface (Table 1). The side chains of these residues can move around freely and are often involved in salt bridge breakage or formation.

Other mutations

In the analysis of all mutations we observed six classes of mutational behaviour:

1. There is only one way of placing the residue without atomic clashes. In all 4 examples in this class we find only one rotamer that leaves the rest of the molecule unaltered and this is the observed rotamer.
2. The conformation of the mutated residue corresponds to the most populated position-specific rotamer. This occurs in approximately 50% of all cases.
3. The most populated rotamer would lead to atomic clashes and the conformation of the mutated residue corresponds to the second most populated rotamer (4 examples).
4. Neither the most populated, nor the second most populated rotamer fits, and the conformation of the mutated residue is very close to the most populated rotamer (2 examples found, both deviate about 20 degrees from the optimal chi1). Never did we observe that the third, i.e., least populated rotamer was selected.
5. The chi1 of the mutated residue corresponds to the most populated rotamer, but the chi2 or chi3 angle is rotated away from the optimal position- specific rotamer distribution to optimise hydrogen bonding. This is observed in almost 15% of all cases (10 examples).
6. There is no preferred position-specific rotamer, or there are insufficient examples in the database, but analysis of potential hydrogen bonds leads to one clear possibility for the structure. This

occurs in 7% of all cases (4 examples). Although in these cases it is possible to fix the charged end of the side-chain in the proper position, it is often difficult to accurately predict the correct conformation of the hydrophobic part of the side-chain.

We have converted these six observed classes into a set of rules. In figure 1 we show a scheme for the prediction of the structure of mutated residues based on these rules. If we assume that the present dataset (see tables 1 A,B) is representative of the universe of mutated residues, then this scheme allows us to correctly predict the local structures of at least 85% of all mutated residues. The actual predictive ability on proteins not studied here may, of course, be lower.

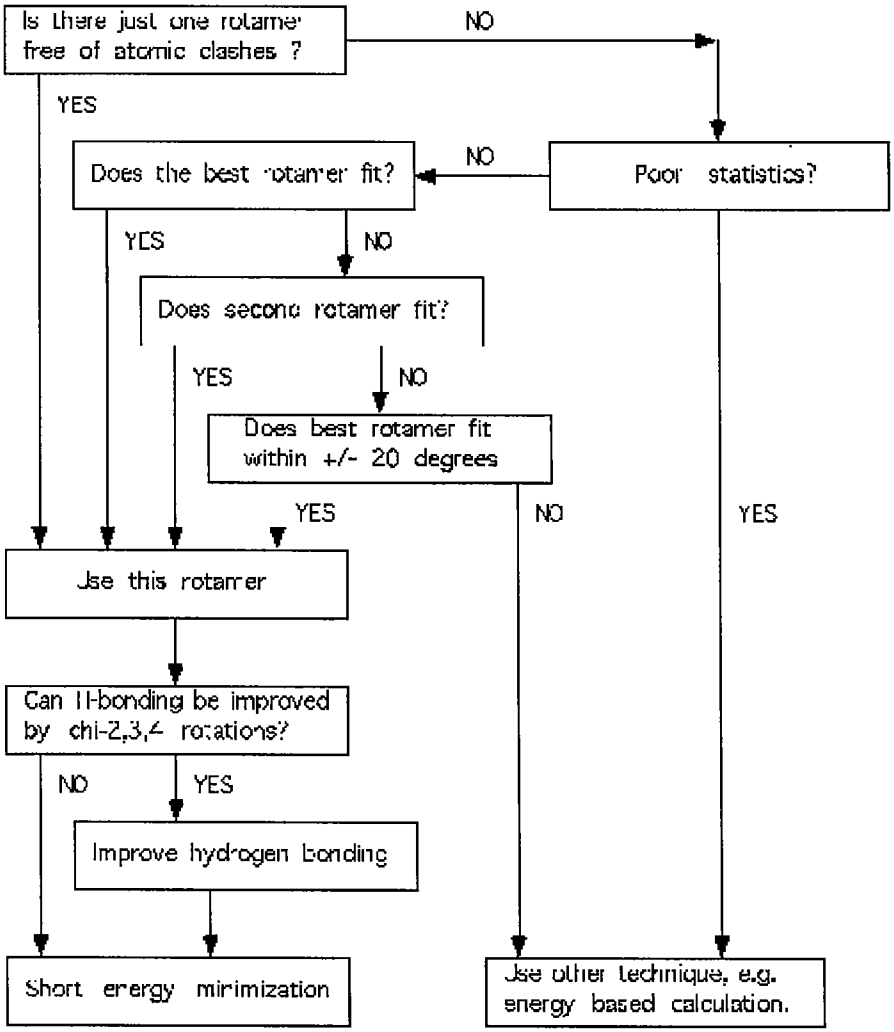

Figure 1. Decision schema for the prediction of point mutant structures.

Of the six rules, the first three are the most clear-cut, i.e. have a high prediction accuracy. These rules cover about 60% of all cases. In all other cases, the prediction accuracy is lower on average. Reassuringly, however, there is a qualitative correlation between correct prediction and the quality of fit of database fragments on the actual backbone, the number of observed database fragments and the sharpness of the chi1 distribution of the central residues in these fragments.

Protein structures are highly optimised for functioning in their parent organism. They are not optimised for thermal stability as can be judged from the relative ease by which proteins can be thermo-stabilised by mutagenesis (Vriend and Eijsink, 1993), and the marginal stability observed in protein (un-)folding experiments. However, the packing of proteins is optimised in the sense that only few residues are not in

an energetically favourable rotameric state. Consequently, core mutations aimed at stabilisation of a protein by means of improved packing are normally unsuccessful (e.g. Vriend and Eijsink, 1993). It seems very costly to disturb the existing delicate balance of forces in the core of a natural protein. Therefore, a mutated residue appears to adapt to its environment rather than the environment to the mutant. Mutations of a big residue into alanine or glycine rarely lead to major rearrangements in the direct 3D environment. This supports the hypothesis that most residues, even in closely packed arrangements, are free of torsional strain, and do not need to use the extra space offered to them to relax. Our results strongly suggest that mutations aimed at improving the packing of the core of a protein will only be successful if the mutated residue can be modelled without atomic clashes in its most preferred position-specific rotamer.

Our dataset cannot be representative of all engineered point mutations. Despite tremendous progress in all aspects of crystallography over the last couple of years, it still is a major endeavour to solve the structure of a mutated protein. The point mutant structures available in the PDB (Bernstein et al., 1972) are therefore not an ideal set of random mutations, but reflect a process of thought about which mutants to make to answer biologically relevant questions. Mutations that modify the core of a protein, and helix capping residues are over represented in table 1. Also, mutations that lead to isomorphous crystals are probably over-represented; this may imply that the percentage of possible point mutations that lead to a large conformational change is somewhat larger than currently apparent.

Not all conformational changes resulting from point mutations can be predicted correctly with the techniques available to date. Examples of changes difficult to predict are large domain motions or local structure adaptations induced by co-factor binding. However, the surprising ease with which the structure of the vast majority of all mutated residue rotamers can be predicted correctly, opens new avenues for planning and analysing protein engineering experiments. Computer screening of large numbers of point mutations becomes feasible, which can save much time and effort otherwise spent on experimental trial and error.

Acknowledgements

We appreciate the support of Angelo Fontana whose continuous encouragement and interest for computational applications in the field of protein stability has contributed much to this project. We thank Olmetta De Filippis for her encouraging support, and for the many helpful discussions. We would like to thank Brian Matthews and his group for providing more than half of our data, and for making data available prior to publication.

References

Abagyan, R.A., Totrov, M.M., J. Mol. Biol. In press.

Alber, T., Dao-pin, S., Wilson, K., Wozniak, J.A., Cook, S.P. and Matthews, B.W. (1987) Nature 330, 41-46.

Alber, T., Bell, J.A., Dao-pin, S., Nicholson, H., Wozniak, J.A., Cook, S. and

Matthews, B.W. (1988) Science 239, 631-635.

Alexander, R.S., Nair, S.K. and Christianson, D.W. Biochemistry 30, 11064-11072.

Baldwin, E.P., Hajiseyedjavadi, O., Baase, W.A. and Matthews, B.W. (1993)Science 262, 1715-1718.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1972)J. Mol. Biol. 112, 535-542.

Bone, R., Shenvi, A.B., Kettner, C.A. and Agard, D.A. (1987)Biochemistry 26, 7609-7614.

Buckle, A.M., Henrick, K. and Fersht, A.R. (1993) J.Mol.Biol. 234, 847-860.

Dao-pin, S., Baase, W.A. and Matthews, B.W. (1990)Prot.: Struct. Funct. Genet. 7, 198-204.

Dao-pin, S., Sauer, U., Nicholson, H. and Matthews, B.W. (1991) Biochemistry 30, 7142-7153.

Dao-pin, S., Alber, T., Baase, W.A., Wozniak, J.A. and Matthews, B.W. (1991)J. Mol. Biol. 221, 647-667.

Dao-pin, S., Anderson, D.E., Baase, W.A., Dahlquist, F.W. and Matthews, B.W. (1991) Biochemistry 30, 11521-11529.

Dao-pin, S., Soderlind, E., Baase, W.A., Wozniak, J.A., Sauer, U. and Matthews, B.W. J. (1991) Mol. Biol. 1, 873-887.

Dunbrack, R.L. and Karplus, M. (1993)J. Mol. Biol. 230, 543-574.

Dunbrack, R.L. and Karplus, M. (1994) Nature Struct. Biol. 1, 334-339.

Finzel, B.C., Clancy, L.L., Holland, D.R., Muchmore, S.W., Watenpaugh, K.D. and

Heinspahr, H.M. (1989) J. Mol. Biol. 209, 779-791.

Folker, P.J.M., Clore, G.M., Driscoll, P.C., Dodt, J., Kohler, S. and Gronenborn, A.M. (1989)Biochemistry 28, 2601-2617.

Grenzin, J., Puras-Lutzke, R., Landt, O., Grunert, H.P., Heinemann, U., Lee,

Saenger, W. and Hahn, U. (1992) J. Mol. Biol. 225, 533-542.

Gilliland, G.L. and Quiocho, F.A. (1981)J. Mol. Biol. 146, 341-362.

Gunsteren, W.F. van and Mark, A.E. (1992)J. Mol. Biol. 227, 389-395.

Grutter, M.G., Hawkes, R.B. & Matthews, B.W.(1979) Nature 277, 667-669.

Gray, T.M. and Matthews, B.W. (1987) J. Biol. Chem. 262, 16858-16864.

Grutter, M.G., Gray, T.M., Weaver, L.H., Alber, T., Wilson, K. and Matthews, B.W. (1987)J. Mol. Biol. 197, 315-329.

Heinz, D.W., Baase, W.A. and Matthews, B.W. (1992) Proc. Natl. Acad. Sci. USA 89, 3751-3755.

Harata, K., Muraki, M., Hayashi, Y. and Jigami, Y. (1992) Protein Sci. 1, 1447-1453.

Hobohm, U., Scharf, M., Schneider, R. and Sander, C. (1992)Prot. Sci. 1, 409-417.

Hurley, J.H., Dean, A.M., Sohl, J.L., Koshland, Jr., D.E. and Stroud, R.M. (1990)Science 249, 1012-1016.

Jones, T.A. and Thirup, S. (1986) EMBO 5 819-823.

Kitamura, S. and Sturtevant, J.M. (1989) Biochemistry 28, 3788-3792 .

Kabsch, W. and Sander, C. (1983)Biopolymers 22, 2577-2637.

Krebs, J.F., Fierke, C.A., Alexander, R.S. and Christianson, D.W. Biochemistry 30, 9153-9160.

Laughton, C.A. (1994) J. Mol. Biol. 235, 1088-1097.

Lewendon, A., Murray, I.A., Shaw, W.V., Gibbs, M.R. and Leslie, A.G.W. (1990)Biochemistry 29, 2075-2080.

Lee, C., Levitt, M. (1991) Nature 352, 448-451.

Loll, P.J.and Lattman, E.E. (1990) Biochemistry 29, 6866-6873.

McGregor, M.J., Islam, S.A., Sternberg, M.J.E., (1987) J. Mol. Biol. 198, 295-310.

Morris, A.L., MacArthur, M.W., Hutchinson, E.G., Thorton, J.M., (1992) PROTEINS 12, 345-364.

Murray, I.A., Lewendon, A., Williams, J.A., Cullis, P.M. and Shaw, W.V. (1991)Biochemistry 30, 3763-3770.

Matsumura, M. and Matthews, B.W. (1989)Science 243, 792-794.

Matsumura, M., Becktel, W.J. and Matthews, B.W. (1988)Nature 334, 406-410.

Matsumura, M., Wozniak, J.A., Dao-pin, S. and Matthews, B.W. (1989)J. Biol.

Chem. 264, 16059-16066.

Muraki, M., Harata, K. and Jigami, Y. (1992)Biochemistry 31, 9212-9219.

Nair, S.K., Calderone, T.L. and Christianson, D.W. (1991) J. Biol. Chem. 266,17320-17325.

Nicholson, H., Becktel, W.J. and Matthews, B.W. (1988) Nature 336, 651-656.

Nicholson, H., Anderson, D.E., Dao-pin, S. and Matthews, B.W. (1991)Biochemistry 30, 9816-9828.

Nicholson, H., Soderlind, E., Tronrud, D.E. and Matthews, B.W. (1989) J. Mol. Biol. 210, 181-193.

Ponder, J.W., Richards, F.M., (1987) J. Mol. Biol. 193, 775-791.

Raag, R., Martinis, S.A., Sligar, S.G. and Poulos, T.L. (1991) Biochemistry 30, 1142 11429.

Simonson, T. and Brunger, A.T. (1992)Biochemistry 31, 8661-8674.

Stites, W.E., Gittis, A.G., Lattman, E.E. and Shortle, D. (1991)J. Mol. Biol. 221,

Stout, C.D. (1989)J. Mol. Biol. 205, 545-555.

Strop, P., Sedlacek, J., Stys, J., Kaderabkova, Z., Blaha, I., Pavlickova, L., Pohl,

Fabry, M., Kostka, V., Newman, M., Frazao, C., Shearer, A., Tickle, I.J. and Blundell, T.L. (1990)Biochemistry 29, 9863-9871.

Thunissen, M.M.G.M., Franken, P.A., Haas, G.H. de, Drenth, J., Kalk, K.H., Tong, L., de Vos, A.M., Milburn, M.V. and Kim, S.H. (1991). J. Mol. Biol. 217, 503-516.

Verheij, H.M., and Dijkstra, B.W. (1993) J. Mol. Biol. 232, 839-855.

Vriend, G. (1990) J. Mol. Graph. 8, 52-56.

Vriend, G. and Eijsink, V.G.H. (1993) J. Comp. Aid. Mol. Des. 7, 367-396.

Vriend, G., Sander, C. and Stouten, P.W.F. (1994) Prot. Eng. 7, 23-29.

Weaver, L.H., Gray, T.M., Grutter, M.G., Anderson, D.E., Wozniak, J.A., Dahlquist, F.W. and Matthews, B.W.(1989) Biochemistry 28, 3793-3797.

Wlodawer, A., Walter, J., Huber, R. and Sjolin, L. (1984)J. Mol. Biol. 180, 301-329.

Zhang, X.J., Baase, W.A. and Matthews, B.W. Biochemistry 30, 2012-2017.

TABLE 1 Mutants used in this study

TABLE 1A: General mutants

MUTATION	sg	el	ei	nRP	RPD chi1	RDH chi2 chi3 chi4	dssp	access	Pred
1L17,Ile3Val	0.1783	0.33	0.28	2	2:1		H	3.67	
1L18,Ile3Tyr	0.3355	0.84	0.65	3	4:2:1	5:1	H	3.67	
1L19,Ser38Asp	0.2035	0.40	0.14	3	4:2:1	3:1		11.53	
1L42,Lys16Glu	0.3053	**1.54**	0.15	-	-		E	15.03	
1L43,Lys16Glu	0.2781	**1.54**	0.08	-	-		E	15.03	
1L34,Arg96His	0.2084	0.35	0.21	2	3:1	4:1	H	3.49	
1L48,Ala98Val	0.3231	0.60	0.46	1	-		H	3.49	
1L54,Met102Lys	0.2663	**2.87**	0.15	2	3:1	3:1 3:1 2:1	H	2.97	
1L57,Asn116Asp	0.4492	0.25	0.28	3	3:2:1	2:1	H	9.26	
1L44,Arg119Glu	0.2767	0.25	0.13	2	2:1	2:1 5:1	H	11.01	
1L45,Lys135Glu	0.1298	0.24	0.10	-		D	S	14.68	
1L20,Asn144Asp	0.2581	0.09	0.17	3	2:1:1	2:1	H	16.42	
1L46,Lys147Glu	0.2720	0.19	0.13	2	3:1	3:1 4:1	H	17.71	
1L53,Val149Cys	0.4980	0.60	0.31	2	3:1		H	0.17	
1L52,Thr152Ser	0.2311	0.29	0.10	3	3:1:1		H	0.52	
1L47,Arg154Glu	0.3055	**4.18**	0.12	2	2:1	2:1 3:1	H	5.77	
1L16,Gly156Asp	0.5067	0.80	0.63	-	D		S	4.54	
1L04,Thr157Asp	0.4789	0.66	0.34	3	3:2:1	2:1	S	9.44	
1L05,Thr157Asp	0.4621	0.66	0.28	3	4:2:1	4:1"	S	9.44	
1L06,Thr157Glu	0.4811	0.28	0.14	-	D		S	9.44	
1L07,Thr157Phe	0.4821	0.47	0.16	3	3:2:1	4:1	S	9.44	
1L09,Thr157His	0.4772	0.50	0.24	3	3:2:1	5:1"	S	9.44	
1L10,Thr157Ile	0.4800	**1.43**	0.25	2	2:1	4:1	S	9.44	

1L11,Thr157Leu	0.4564	0.76	0.21	2	5:1	5:1	S	9.44
1L12,Thr157Asn	0.4601	0.29	0.31	3	3:2:1	2:1	S	9.44
1L13,Thr157Arg	0.4972	0.23	0.10	1	1	1 1 1	S	9.44
1L14,Thr157Ser	0.4743	0.34	0.07	3	4:2:1		S	9.44
1L15,Thr157Val	0.4791	0.80	0.20	2	2:1		S	9.44
1L27,Pro86Asp	0.1286	0.90	0.19	2	5:1	4:1	H	12.23
1L29,Pro86His	0.2361	**1.41**	0.24	2	3:2	2:1	H	12.23
1L30,Pro86Leu	0.2384	0.90	0.24	2	4:1	3:1	H	12.23
1L31,Pro86Arg	0.4824	**1.44**	0.38	-	D		H	12.23
1L32,Pro86Ser	0.2516	**1.54**	0.24	-	D		H	12.23
1L55,Asp92Asn	0.4724	0.54	0.26	2	4:1	3:1		13.11
1L61,Ser38Asn	0.3882	0.48	0.16	3	5:2:1	3:1		11.53
1L62,Thr109Asp	0.4086	0.66	0.12	3	4:2:1	3:1	H	20.27
1L59,Thr109Asn	0.4110	0.56	0.06	3	2:1:1	2:1	H	20.27
1L37,Thr115Glu	0.2233	0.66	0.22	-	D		H	16.25
1L38,Gln123Glu	0.2985	**2.51**	0.05	2	3:1	4:1	T 3:1	8.74
1L39,Asn144Glu	0.2586	0.53	0.24	2	3:1	3:2:1	H 3:1"	16.43
1L40,Asn144Glu	0.2561	0.48	0.24	2	2:1	4:2:1	H 3:1	16.43
1LAA,Asp53Glu	0.5071	**4.26**	0.16	3	4:2:1	3:1	E 2:1	3.14
1TBY,Tyr63Leu	0.5116	**1.46**	0.17	1		2:1	T	9.96
1TCY,Tyr63Phe	0.4892	0.74	0.06	3	8:2:1	5:1	T	9.96
1TDY,Tyr63Trp	0.4798	0.76	0.14	3	4:2:1	3:1	T	9.96
5CA2,Thr200Ser	0.3491	**2.62**	0.40	3	3:2:1		S	6.29
6CA2,Val143Phe	0.3736	**1.23**	0.27	3	4:2:1	4:1	E	2.1
8CA2,Val143His	0.3458	0.73	0.29	3	5:2:1	2:1	E	2.1
9CA2,Val143Tyr	0.3603	**1.07**	0.28	3	3:1:1	4:1	E	2.1
1HEA,Leu198Arg	0.3583	0.60	0.18	-	D			8.97
1HEC,Leu198His	0.2777	0.86	0.31	-	D			8.97
6ICD,Ser113Asp	0.2461	0.80	0.09	3	8:1:1	5:1		11.7
7ICD,Ser113Glu	0.2421	0.52	0.28	-	D			11.7
4CLA,Leu160Phe	0.2140	**1.05**	0.09	3	4:2:1	5:1		7.5
2CLA,Asp199Asn	0.4006*	**5.62**	0.66	2	4:1	5:1		9.3
2Q21,Gly12Val	0.4915	**1.22**	0.13	2	3:1		T	19.4
3CMS,Val111Phe	0.7926	**2.45**	0.18	3	4:2:1	4:1	H	4.4
2SNM,Val66Lys	0.5846	**2.44**	0.14	2	3:2	2:1"	H	2.8
6ABP,Met108Leu	0.4185	0.60	0.40	2	3:1	2:1	E	5.2

TABLE 1B: X -> Ala and X -> Gly mutants

MUTATION	sg	el	e1	dssp	access
1L33,Val131Ala	0.2272	0.66	0.19	H	14.15
1L69,Leu133Ala	0.4243	0.85	0.27	H	3.14
1L02,Thr157Ala	0.4647	0.33	0.15	S	9.44
1L58,Pro143Ala	0.2379	0.42	0.44	H	8.34
1L21,Asn55Gly	0.2318	0.30	0.38	T	16.95
1L08,Thr157Gly	0.4604	0.39	0.23	S	9.44
1L22,Lys124Gly	0.2180	0.40	0.23	T	11.53
1L66,Lys43Ala	0.4466	0.42	0.16	H	8.91
1L67,Leu46Ala	0.1750	0.39	0.20	H	0.52
1L68,Ser44Ala	0.4475	0.56	0.25	H	14.50
1L65,Asp47Ala	0.4833	**1.56**	0.08	H	7.70
1L25,Pro86Ala	0.2686	**1.62**	0.21	H	12.23

1L28,Pro86Gly	0.1747	**1.35**	0.22	H	"
1P08,Met192Ala	0.1459	**1.70**	0.08		1.2
1P09,Met213Ala	0.3160	**4.49**	0.14	E	0.0
21BI,Cys71Ala	0.6015	0.86	0.20	E	0.2
41BI,Cys8Ala	e	**3.37**	0.08	E	0.0
2FD2,Cys24Ala	0.2313	0.60	0.45		0.2
1CLA,Ser148Ala	0.1547	**1.37**	0.09	E	8.0
2CP4,Thr252Ala	0.3763	0.90	0.70	T	8.4
1TAY,Tyr63Ala	0.4863	**1.58**	0.23	T	9.9
12CA,Val121Ala	0.2671	0.63	0.09	E	3.5
7CA2,Val143Gly	0.3632	0.52	0.48	E	2.1
1APB,Pro254Gly	0.4113	0.54	0.57	E	0.5

sg : RMS deviation of all atoms after superposing wild type and mutated protein

el : Largest atomic movement in the environment of the mutated residue. Changes > 1.0 Angstrom are in bold.

ei : Largest atomic movement in the backbone of the mutated residue.

nRP : Number of observed distinct populations in chi1.

RPD : Relative frequency of occurrence in the nRP populations in chi1.

RDH : Relative frequency of occurrence in distinct populations of chi2, chi3, chi4 (only given if applicable).

dssp : Secondary structure classification according to DSSP (Kabsch and Sander, 1983).

access : Accessible surface area of beta position for the amino acid side-chain, a mutating it into an alanine.

yes : Predicted correctly using the most populated rotamer.

no : Misprediction.

C : Poor statistics. Structure prediction often correct, but poorly reproducible. In case of long charged side chains the polar atoms are normally placed more accurately than the hydrophobic ones.

D : Very diffuse distribution.

a : The higher c was modified to improve hydrogen bonding.

b : chi1 differs up to 20 degrees from the preferred position-specific rotamer

c : Side-chain in second best position-specific rotamer because of bump in best rotamer.

d : There is only one way to accommodate the new side-chain without bumps.

e : Too many side chain atoms missing in mutant structure to determine RMS deviation.

Mutants used in this study (Stites et al., 1991; Strop et al., 1990; Raag et al., 1991; Wlodawer et al., 1984; Bone et al., 19487; Gilliland and Quiocho, 1981; Loll and Lattman, 1990; Matsumura and Matthews, 1989; Matsumura et al., 1988; Matsumura et al., 1989; Nicholson et al., 1988; Nicholson et al., 1991; Dao-pin et al., 1990; Nicholson et al., 1989; Dao-pin et al., 1991; Alber et al., 1988; Grutter et al., 1979; Weaver et al., 1989; Kitamura and Sturtevant, 1989; Dao-pin et al., 1991; Dao- pin et al., 1991; Zhang et al., 1991; Alber et al., 19487; Gray and Matthews, 19487; Grutter et al., 19487; Heinz et al., 1992; Harata et al., 1992; Muraki et al., 1992; Krebs et al., 1991; Alexander et al., 1991; Nair et al., 1991; Finzel et al., 1989; Hurley et al., 1990; Lewendon et al., 1990; Murray et al., 1991)