FOR THE RECORD

Change in dimerization mode by removal of a single unsatisfied polar residue located at the interface

P.R. POKKULURI, X. CAI, G. JOHNSON, F.J. STEVENS, AND M. SCHIFFER

Argonne National Laboratory, Biosciences Division, Argonne, Illinois 60439

(RECEIVED March 20, 2000; FINAL REVISION May 31, 2000; ACCEPTED June 2, 2000)

Abstract: The importance of unsatisfied hydrogen bonding potential on protein–protein interaction was studied. Two alternate modes of dimerization (conventional and flipped form) of an immunoglobulin light chain variable domain (V_L) were previously identified. In the flipped form, interface residue Gln89 would have an unsatisfied hydrogen bonding potential. Removal of this Gln should render the flipped dimer as the more favorable quaternary form. High resolution crystallographic studies of the Q89A and Q89L mutants show, as we predicted, that these proteins indeed form flipped dimers with very similar interfaces. A small cavity is present in the Q89A mutant that is reflected in the ~ 100 times lower association constant than found for the Q89L mutant. The association constant of Q89A and Q89L proteins ($4 \times 10^6 \text{ M}^{-1}$ and $>10^8 \text{ M}^{-1}$) are 10- and 1,000-fold higher than that of the wild-type protein that forms conventional dimers clearly showing the energetic reasons for the flipped dimer formation.

Keywords: antibody light chain; association constant; crystal structure; dimer interface; quaternary interactions

Unsatisfied polar side chains buried in the hydrophobic interior of proteins are known to reduce the stability of the protein (Hendsch et al., 1996). In case of dimer formation, such a buried polar side chain would be expected to diminish the dimerization constant by reducing the total interaction energy between the monomers. Unsatisfied hydrogen bonding potential is very rare in the interior of proteins, $<1\%$ of side chain donors, and $<3\%$ of side-chain acceptors are not hydrogen bonded (McDonald & Thornton, 1994). Only about 2% of NE2 or OE1 atoms of buried glutamine residues were found not to form hydrogen bonds in high resolution protein structures (McDonald & Thornton, 1994). Protein–protein interfaces are not so well optimized in terms of satisfying the hydrogen bonding potential as 17.4% of fully buried polar residues were found not to form hydrogen bonds or to have compensating charges nearby (Xu et al., 1997).

We previously found (Pokkuluri et al., 1998) that immunoglobulin light chain variable domains (V_L) dimerize in two different arrangements using the same faces of the domains: (1) The conventional form in which the V_L domains are paired in parallel as found in Fab fragments; and (2) the flipped form, in which the V_L domains are paired in an antiparallel orientation as formed by myelin P₀ adhesion domains (Shapiro et al., 1996). In conventional V_L domain dimers (Epp et al., 1975; Huang et al., 1994, 1997), Gln89 is partially exposed to the solvent and it interacts with Tyr36 of the same monomer and with Tyr36 or Gln89 across the dimer interface. In the flipped dimers (Pokkuluri et al., 1998), in the Q38E and K30T mutants of human $\kappa 4$ light chain Len, both Tyr36 and Gln89 are buried. The hydrogen bonding capacity of Gln89 is only partially satisfied. To eliminate the unsatisfied hydrogen bonding potential in the flipped conformation, and thereby favor its formation, residue Gln89 was replaced by hydrophobic residues Leu and Ala and the structures of these mutants were determined by X-ray diffraction.

Results and discussion: The structures of the Q89A and Q89L mutants were determined by molecular replacement. The R -factors are 0.186 and 0.193 for 2 and 1.8 Å data, respectively. Data collection and refinement statistics are shown in Table 1. Ramachandran plots for both structures indicated that 90% of the residues were in the most favored regions, only Ala51 was in the disallowed region as observed previously by Steipe et al. (1992). The Q89L and Q89A mutants formed flipped dimers; their mode of dimerization is similar to what was observed for the Q38E mutant. The root-mean-square deviation (RMSD) between Q89A and Q89L dimers is 1.57 Å. Upon superimposing one of the monomers of Q89A and Q89L 0.3 Å translation and 13.4° rotation are required to superimpose the second monomer. These values are 1.2 Å and 4.8° for Q38E and Q89A superposition, and 1.2 Å and 7.5° for Q38E and Q89L superposition. The surface areas buried in the dimers are 1,746 Å² and 1,561 Å², for Q89A and Q89L, respectively. These values are similar to 1,736 Å² found in the Q38E mutant, but differ significantly from the 1,150 Å² of buried surface area of the conventional dimer of native V_L .

The number of interdomain hydrogen bonds, ten and six in Q89A and Q89L, respectively, increased compared to the native

Reprint requests to: Marianne Schiffer, Biosciences Division, Argonne National Laboratory, Argonne, Illinois 60439; e-mail: mschiffer@anl.gov.

Table 1. Data collection and refinement statistics^a

	Q89A	Q89L
Space group	P6 ₂ 22	P2 ₁
<i>a</i> (Å)	106.3	35.3
<i>b</i> (Å)	106.3	83.6
<i>c</i> (Å)	45.1	42.8
γ or β (deg)	120.0	109.0
<i>V_M</i> (Å ³ /Da)	3.0	2.5
<i>R</i> merge (%) (last shell)	7.4 (24.2)	5.6 (17.2)
Highest resolution (Å)	2.0	1.8
No. of reflections used (<i>F_o</i> > 3σ <i>F_o</i>)	9,087	19,237
Final <i>R</i> -factor	0.186	0.193
Free <i>R</i> -factor (10% data)	0.218	0.226
No. of refined protein atoms	883	1,763
No. of solvent molecules ^a	158 + 1Zn ²⁺	305
RMSD bond length (Å)	0.005	0.006
RMSD bond angle (deg)	1.4	1.4
Mean temperature factor		
For protein atoms (Å ²)	16.1	15.2
For solvent atoms (Å ²)	31.5	26.3

^aThe solvent molecules for Q89A include a Zn²⁺ ion positioned on the twofold parallel to the *z*-axis. The nitrogens, carbonyl oxygens, and side chains of the twofold related N-terminal aspartic acid residues form a distorted octahedron around the Zn ion.

dimer. These numbers are comparable to the ones found in Q38E and K30T (Pokkuluri et al., 1998). As was previously observed for Glu38 in the Q38E mutant, in both Q89A and Q89L structures Gln38 forms a hydrogen bond with Tyr96 of the other monomer. The hydrogen bonds between Tyr87 and Tyr96 are also maintained. In the Q89A mutant, both the hydrogen bonds between Thr94 and Gln100 and the salt bridge between Lys45 and Glu55

are present. Though the above residues are spatially close in the Q89L structure, the electron density is poor for residues Lys45 and Gln100 in both monomers, therefore the hydrogen bonds are present only part of the time.

In the Q89L structure, Tyr36 interacts through a water molecule with the backbone carbonyl oxygen of proline 44 of the other monomer as was observed in Q38E and K30T mutants. In those mutants, the (tentatively identified) OE1 oxygen of Gln89 accepts a hydrogen bond from Tyr36 of the same monomer but the NE2 atom does not have suitable neighbors to form hydrogen bonds. The closest residues to it are Phe98 and Tyr96 from the same domain and Phe98 from the twofold-axis related domain. In the Q89L structure, the interactions of Leu89 with the neighboring residues are favorable; it forms hydrophobic contacts with residue Tyr96 of the same domain and residues Phe98 of both domains (Fig. 1). Ala89 is in van der Waals contact with atoms of the phenyl rings of Phe98 and Tyr96 of the same monomer.

To minimize the cavity at the dimer interface created by the Ala substitution in the Q89A mutant, a minor rearrangement of the dimer interface occurs compared to Q89L. The cavity near Ala89 of each monomer has a volume of 61 Å³ and contains no water molecules. Without the domain rearrangement, the cavity would have a volume of 77 Å³. As a result of the rearrangement, Tyr36 OH in the Q89A mutant forms a direct hydrogen bond (2.71 Å) with the carbonyl oxygen of residue 44. The α -carbon atoms of residues 36 of the dimer are 1 Å closer in Q89A (12.3 Å) than in Q89L (13.8 Å) structure. The distance between α -carbons of residues 89 is the same in both structures.

The importance of buried polar residues on determining the energetics of dimer formation is illustrated by the properties of mutants Q89L and Q89A. The association constants (*K_a*) of Q89L and Q89A are >10⁸ M⁻¹ and 4 × 10⁶ M⁻¹, respectively. The association constants of both the Q89L and Q89A mutants increased significantly compared to the wild-type conventional dimer (*K_a* 2 × 10⁵ M⁻¹) and flipped dimers Q38E and K30T

Fig. 1. The close packing of hydrophobic residues Leu89, Tyr96, Phe98, at the center of the Q89L dimer is shown in the top figure. The lower figure shows the interface of the Q89A dimer. In addition to the difference in domain packing between Q89L and Q89A, as discussed in the text, the orientation of side chain Phe98 also altered. The figure was generated by SETOR (Evans, 1990).

(K_a $5.8 \times 10^5 \text{ M}^{-1}$ and $0.8 \times 10^5 \text{ M}^{-1}$). The hundred times higher association constant of the Leu mutant compared to the Ala mutant is probably due to increased hydrophobic interactions combined with a large favorable desolvation component of the leucine residue compared to that of the alanine. The cavity created at the interface of the Ala mutant is also likely to diminish its dimerization.

We have exploited the remarkable dimerization properties of the V_L domain of Len. Though the formation of the flipped dimer is favored by a larger buried surface area, and by the presence of more salt bridges and hydrogen bonds than exist in the conventional form, it does not occur in the wild-type protein. We found that its formation is prevented by a second independent unfavorable factor in addition to the one previously identified. An unsatisfied polar residue (Q89) that is buried in the interface. Glutamine residue at position 89 favors the conventional over the flipped dimer, since in the flipped conformation it is buried, without all of its hydrogen bonding potential satisfied. This impediment of flipped dimer formation is removed by the Q89L or Q89A mutations, which substitute a hydrophobic residue at position 89. As we discussed previously, excess positive potential at the interface can prevent flipped dimer formation (Pokkuluri et al., 1998). The excess positive potential can be removed by either the addition of a negatively charged residue (Q38E) or the removal of a positively charged residue (K30T) near the interface. In each mutated protein Q89L, Q89A, Q38E, and K30T, only one of the unfavorable factors that prevent the flipped dimer formation is removed. The Q89L and Q89A mutations eliminate the polar residues at the interface but not the excess positive potential. The Q38E and K30T mutations remove the excess positive potential but not the buried polar residue. The association constants reflect both the positive and negative influences of the above residues, they suggest that the removal of the unsatisfied hydrogen bonding capacity is energetically more favorable for the flipped dimer formation than the removal of the excess positive potential from the interface.

Conclusions: In this work, we have shown that the mode of dimerization of a V_L β -domain can be changed from the conventional form to the flipped form by substituting a hydrophobic residue for a polar residue at the interface. The substitution of the larger hydrophobic residue, Leu, for Gln89 not only made the flipped dimer formation favorable, but also significantly increased the dimerization constant.

In this system, we demonstrated that rules derived from amino acid substitutions in protein cores, as expected, can be extended to interactions that take place during formation of protein-protein complexes. The dimer formation was influenced by the negative effect of unsatisfied polar groups, the benefit of desolvation of nonpolar groups, and the impaired stability resulting from packing flaws. Thus, the quantitative analyses of amino acid variations at dimerization interfaces could provide a complementary approach for understanding the processes of side-chain packing in the protein core.

Materials and methods: LenQ89A and LenQ89L mutants were constructed, expressed, and purified as described earlier by Wilkins Stevens et al. (1995). Diffraction data were collected on an R-Axis IIC image plate detector mounted on a Rigaku RU200 rotating anode (Q89A, Q89L first data set) and at the Structural Biology Center's beamline at the Advanced Photon Source (Q89L second data set). The diffraction data were processed with Denzo, then

merged and scaled using Scalepack (Otwinowski, 1992). The structures were determined by molecular replacement using AMoRe (Navaza, 1994). The structures were refined with the program CNS (Brünger et al., 1998) using a maximum likelihood target for refinement; *B*-factor and bulk solvent corrections were applied. Manual rebuilding was performed using the program CHAIN (Sack, 1988).

LenQ89A was crystallized by hanging drops from 1.8 M ammonium sulfate, 0.1 M sodium HEPES, pH 7.5, 0.001 M zinc acetate. The protein concentration was 10 mg/mL. Before flash cooling of the crystals for data collection, a crystal was briefly soaked in a cryoprotectant solution containing 20% sucrose in addition to all other constituents (except the protein) of the mother liquor in which the crystals were grown and directly introduced in a cold nitrogen stream. The structure was determined by molecular replacement using a monomer of the native Len as the search model (Protein Data Bank (PDB) code, 1LVE; GLN89 side-chain atoms beyond CB were removed).

LenQ89L was crystallized by hanging drops from 18% PEG8000, 0.2 M ammonium sulfate. Before flash cooling for data collection, a crystal was briefly soaked in a cryoprotectant solution containing 28% PEG8000 in addition to all other constituents (except the protein) of the mother liquor in which the crystals were grown. Even though the crystal diffracted to about 2.4 Å, due to problems encountered during data collection the data set was only 75% complete to 2.8 Å. Crystals belonged to spacegroup $P2_12_12_1$ with the unit cell dimensions of $a = 66.1$, $b = 42.7$, $c = 83.5$ Å, and contained one V_L dimer per asymmetric unit ($V_M = 2.5$). The structure was determined by molecular replacement with a dimer of LenQ38E as a search model (PDB code, 3LVE; dimer generated by crystallographic operation). The resulting model was then further refined to an *R*-factor of 24.6% and an *R*-free (Brünger, 1997) of 30.4%. Since we were not able to reproduce the above crystals to get a complete data set, we looked for other conditions for crystallization. Crystals were grown from 18% PEG8000, 0.2 M ammonium sulfate at 4 °C. These crystals were found to belong to $P2_1$ spacegroup with a dimer in the asymmetric unit. The model of Q89L from $P2_12_12_1$ crystal form was used to solve the rotation and translation functions.

Both structures and structure factors were deposited in the Protein Data Bank; ID codes for Q89A, Q89L are 5lve and 1qac, respectively.

Association constants were determined by small-zone column chromatography (Stevens, 1986; Raffin et al., 1997). The cavity volumes were calculated with the program GRASP (Nicholls et al., 1991) using a probe radius of 1.4 Å. The buried surface areas were determined by using the program Surface (Lee & Richards, 1971) with a probe radius of 1.4 Å.

Acknowledgments: This work was supported by the U.S. Department of Energy, the Office of Biological and Environmental Research under contract No. W-31-109-Eng-38 and by the U.S. Public Health Service Grant DK43757. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Energy Research, under contract No. W-31-109-Eng-38.

References

- Brünger AT. 1997. The free *R* value: A more objective statistic for crystallography. *Methods Enzymol* 277:366–396.
- Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve W,

- Jiang J-S, Kuszewski J, Nilges M, Pannu NS, et al. 1998. Crystallography and NMR system: A new software system for macromolecular structure determination. *Acta Crystallogr D* 54:905–921.
- Epp O, Lattman EE, Schiffer M, Huber R, Palm W. 1975. The molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI refined at 2.0 Å resolution. *Biochemistry* 14:4943–4952.
- Evans SV. 1990. SETOR: Hardware lighted three-dimensional solid model representations of macromolecules. *J Mol Graphics* 11:134–138.
- Hendsch ZS, Jonsson T, Sauer RT, Tidor B. 1996. Protein stabilization by removal of unsatisfied polar groups: Computational approaches and experimental tests. *Biochemistry* 35:7621–7625.
- Huang D-B, Chang C-H, Ainsworth C, Brünger AT, Eulitz M, Solomon A, Stevens FJ, Schiffer M. 1994. Comparison of crystal structures of two homologous proteins: Structural origin of altered domain interactions in immunoglobulin light-chain dimers. *Biochemistry* 33:14848–14857.
- Huang D-B, Chang C-H, Ainsworth C, Johnson G, Solomon A, Stevens FJ, Schiffer M. 1997. Variable domain structure of the nonpathological human κ IV light chain Len: High homology to the murine light chain McPC603. *Mol Immunol* 34:1291–1301.
- Lee B, Richards FM. 1971. Solvent accessibility of groups in proteins. *J Mol Biol* 55:379–400.
- McDonald IK, Thornton JM. 1994. Satisfying hydrogen bonding potential in proteins. *J Mol Biol* 238:777–793.
- Navaza J. 1994. AMoRE: An automated package for molecular replacement. *Acta Crystallogr A* 50:157–163.
- Nicholls A, Sharp KA, Honig B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11:281–296.
- Otwinowski Z. 1992. *An oscillation data processing suite for macromolecular crystallography*. New Haven, CT: Yale University.
- Pokkuluri PR, Huang D-B, Raffin R, Cai X, Johnson G, Wilkins Stevens P, Stevens FJ, Schiffer M. 1998. A domain flip as a result of a single amino-acid substitution. *Structure* 6:1067–1073.
- Raffin R, Wilkins Stevens P, Boogard C, Schiffer M, Stevens FJ. 1997. Reengineering immunoglobulin domain interactions by introduction of charged residues. *Protein Eng* 11:303–309.
- Sack JS. 1988. CHAIN—A crystallographic modeling program. *J Mol Graphics* 6:224–225.
- Shapiro L, Doyle JP, Hensley P, Colman DR, Hendrickson WA. 1996. Crystal structure of the extracellular domain from P₀, the major structural protein of peripheral nerve myelin. *Neuron* 17:435–449.
- Steipe B, Plückthun A, Huber R. 1992. Refined crystal structure of a recombinant immunoglobulin domain and a complementarity-determining region 1-grafted mutant. *J Mol Biol* 225:739–753.
- Stevens FJ. 1986. Analysis of protein-protein interaction by simulation of small-zone size-exclusion chromatography: Application to an antibody-antigen association. *Biochemistry* 25:981–993.
- Wilkins Stevens P, Raffin R, Hanson DK, Deng Y-L, Berrios-Hammond M, Westholm FA, Murphy C, Eulitz M, Wetzel R, Solomon A, et al. 1995. Recombinant immunoglobulin variable domains generated from synthetic genes provide a system for in vitro characterization of light-chain amyloid proteins. *Protein Sci* 4:421–432.
- Xu D, Tsai C-J, Nussinov R. 1997. Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Eng* 10:999–1012.